Infrared devices and techniques

A. ROGALSKI[*1] and K. CHRZANOWSKI[2]

[1]Institute of Applied Physics, [2]Institute of Optoelectronics
Military University of Technology, 2 Kaliskiego Str., 00-908 Warsaw, Poland

The main objective of this paper is to produce an applications-oriented review covering infrared techniques and devices. At the beginning infrared systems fundamentals are presented with emphasize on thermal emission, scene radiation and contrast, cooling technics, and optics. Special attention is put on night vision and thermal imaging concepts. Next section shortly concentrates on selected infrared systems and is arranged in order to increase complexity; from smart weapon seekers, image intensifier systems, thermal imaging systems, to space-based systems. Finally, other important infrared techniques and devices are shortly described between them the most important are: non-contact thermometers, radiometers, LIDAR, and gas sensors.

1. Introduction

Looking back over the past 1000 years we notice that infrared (IR) radiation itself was unknown until 200 years ago when Herschel's experiment with thermometer was first reported [1]. He built a crude monochromator that used a thermometer as a detector so that he could measure the distribution of energy in sunlight. Following the works of Kirchhoff, Stefan, Boltzmann, Wien, and Rayleigh, Max Planck culminated the effort with well-known Planck's law.

Traditionally, IR technologies are connected with controlling functions and night vision problems with earlier applications connected simply with detection of IR radiation, and later by forming IR images form temperature and emissivity differences (systems for recognition and surveillance, tank sight systems, anti-tank missiles, air-air missiles). The years during World War II saw the origins of modern IR techniques. Recent success in applying IR technology to remote sensing problems has been made possible by the successful development of high-performance IR detectors over five decades. Most of the funding has been provided to fulfil military needs, but peaceful applications have increased continuously, especially in the last decade of twentieth century. These include medical, industry, earth resources, and energy conservation applications. Medical applications include thermography in which IR scans of the body detect cancers or other trauma, which raise the body surface temperature. Earth resources determinations are done by using IR images from satellites in conjunction with field observation for calibration (in this manner, e.g. the area and content of fields and forests can be determined). In some cases even the state of health of a crop be determined from space. Energy conservation in homes and industry has been aided by the use of IR scans to determine the points of maximum heat loss. Demands to use these technologies are quickly growing due to their effective applications, e.g., in global monitoring of environmental pollution and climate changes, long time prognoses of agriculture crop yield, chemical process monitoring, Fourier transform IR spectroscopy, IR astronomy, car driving, IR imaging in medical diagnostics, and others.

Today, only about 10% of the market is commercial. After a decade the commercial market can grow to over 70% in volume and 40% in value, largely connected with volume production of uncooled imagers for automobile driving [2]. In large volume production for automobile drivers the cost of uncooled imaging systems will decrease to below $1000.

The infrared range covers all electromagnetic radiation longer than the visible, but shorter than millimetre waves. Many proposals of division of IR range have been published. The division shown below is based on limits of spectral bands of commonly used IR detectors. Wavelength 1 μm is a sensitivity limit of popular Si detectors. Similarly, wavelength 3 μm is a long wavelength sensitivity of PbS and InGaAs detectors; wavelength 6 μm is a sensitivity limit of InSb, PbSe, PtSi detectors and HgCdTe detectors optimised for 3–5 μm atmospheric window; and finally wavelength 15 μm is a long wavelength sensitivity limit of HgCdTe detectors optimised for 8–14 μm atmospheric window.

[*]e-mail: rogan@wat.waw.pl

Table 1. Division of infrared radiation.

Region (abbreviation)	Wavelength range (µm)
Near infrared (NIR)	0.78–1
Short wavelength IR (SWIR)	1–3
Medium wavelength IR (MWIR)	3–6
Long wavelength IR (LWIR)	6–15
Very long wavelength IR (VLWIR)	15–1000

2. Infrared systems fundamentals

2.1. Thermal emission

All objects are composed of continually vibrating atoms, with higher energy atoms vibrating more frequently. The vibration of all charged particles, including these atoms, generates electromagnetic waves. The higher the temperature of an object, the faster the vibration, and thus the higher the spectral radiant energy. As a result, all objects are continually emitting radiation at a rate with a wavelength distribution that depends upon the temperature of the object and its spectral emissivity, $\varepsilon(\lambda)$.

Radiant emission is usually treated in terms of the concept of a blackbody. A blackbody is an object that absorbs all incident radiation and, conversely according to the Kirchhoff's law, is a perfect radiator. The energy emitted by a blackbody is the maximum theoretically possible for a given temperature. The radiative power (or number of photon emitted) and its wavelength distribution are given by the Planck radiation law

$$W(\lambda,T) = \frac{2\pi hc^2}{\lambda^5}\left[\exp\left(\frac{hc}{\lambda kT}\right) - 1\right]^{-1} \quad \text{W/(cm}^2\text{ µm)}, \quad (1)$$

$$P(\lambda,T) = \frac{2\pi c}{\lambda^4}\left[\exp\left(\frac{hc}{\lambda kT}\right) - 1\right]^{-1} \quad \text{photons/(s cm}^2\text{ µm)}, \quad (2)$$

where λ is the wavelength, T is the temperature, h is the Planck's constant, c is the velocity of light, and k is the Boltzmann's constant.

Figure 1 shows a plot of these curves for a number of blackbody temperatures. As the temperature increases, the amount of energy emitted at any wavelength increases too, and the wavelength of peak emission decreases. The latter is given by the Wien's displacement law

$$\lambda_{mw}T = 2898 \ \mu mK \quad \text{for maximum watts,}$$

$$\lambda_{mp}T = 3670 \ \mu mK \quad \text{for maximum photons.}$$

The loci of these maxima are shown in Fig. 1. Note that for an object at an ambient temperature of 259 K, λ_{mw} and

λ_{mp} occur at 10.0 µm and 12.7 µm, respectively. We need detectors operating near 10 µm if we expect to "see" room temperature objects such as people, trees and truck without the aid of reflected light. For hotter objects such as engines, maximum emission occurs at shorter wavelengths. Thus, the waveband 2–15 µm in infrared or thermal region of the electromagnetic spectrum contains the maximum radiative emission for thermal imaging purposes.
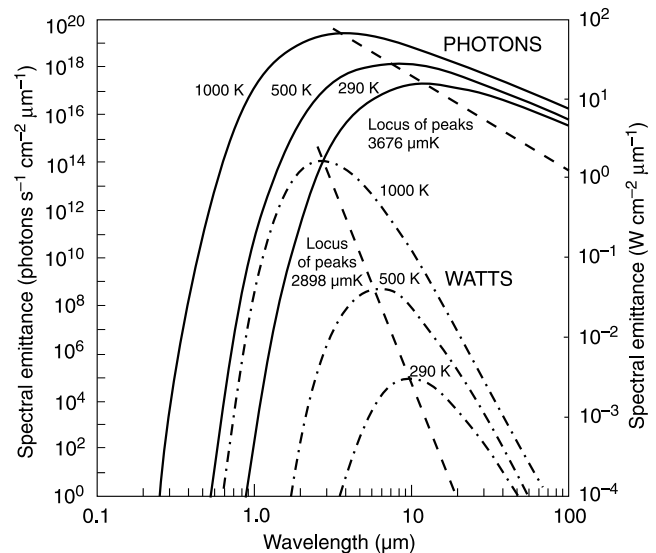

Fig. 1. Planck's law for spectral emittance (after Ref. 3).

2.2. Atmospheric transmission

Most of the above mentioned applications require transmission through air, but the radiation is attenuated by the processes of scattering and absorption. Scattering causes a change in the direction of a radiation beam; it is caused by absorption and subsequent reradiation of energy by suspended particles. For larger particles, scattering is independent of wavelength. However, for small particles, compared with the wavelength of the radiation, the process is known as Rayleigh scattering and exhibits a λ^{-4} dependence. Therefore, scattering by gas molecules is negligibly small for wavelengths longer than 2 µm. Also smoke and light mist particles are usually small with respect to IR wavelengths, and IR radiation can therefore penetrate further through smoke and mists than visible radiation. However, rain, fog particles and aerosols are larger and consequently scatter IR and visible radiation to a similar degree.

Figure 2 is a plot of the transmission through 6000 ft of air as a function of wavelength. Specific absorption bands of water, carbon dioxide and oxygen molecules are indicated which restricts atmospheric transmission to two windows at 3–5 µm and 8–14 µm. Ozone, nitrous oxide, carbon monoxide and methane are less important IR absorbing constituents of the atmosphere.
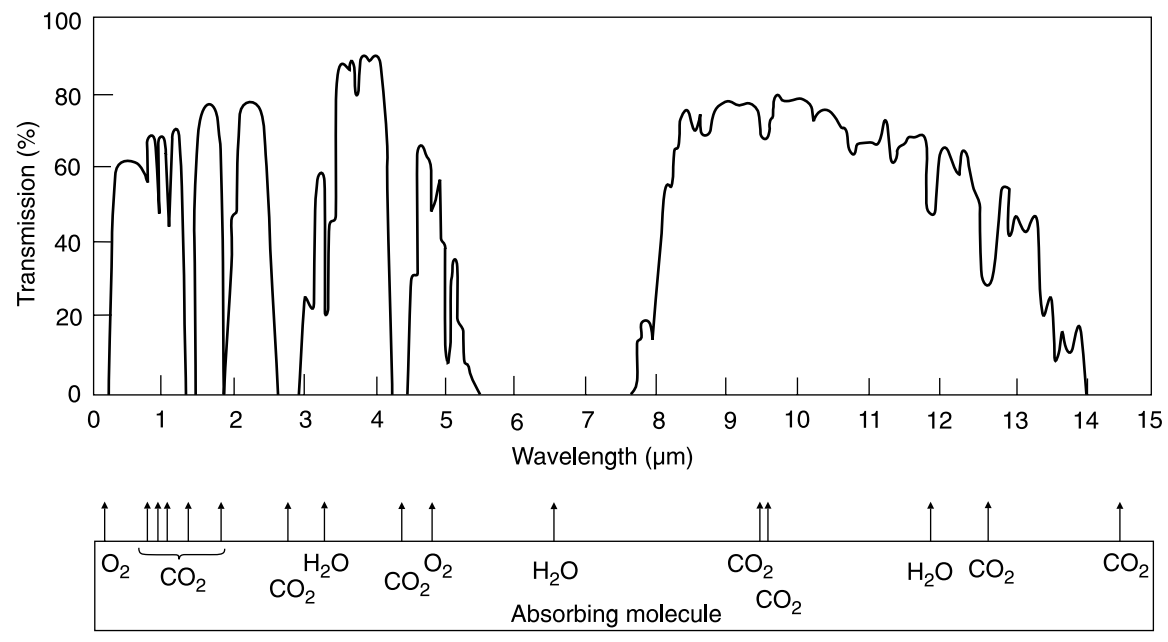
Fig. 2. Transmission of the atmosphere for a 6000 ft horizontal path at sea level containing 17 mm of precipitate water (after Ref. 4).

2.3. Scene radiation and contrast

The total radiation received from any object is the sum of the emitted, reflected and transmitted radiation. Objects that are not blackbodies emit only the fraction $\varepsilon(\lambda)$ of blackbody radiation, and the remaining fraction, $1-\varepsilon(\lambda)$, is either transmitted or, for opaque objects, reflected. When the scene is composed of objects and backgrounds of similar temperatures, reflected radiation tends to reduce the available contrast. However, reflections of hotter or colder objects have a significant effect on the appearance of a thermal scene. The powers of 290 K blackbody emission and ground-level solar radiation in MWIR and LWIR bands are given in Table 2. We can see that while reflected sunlight have a negligible effect on 8–13 μm imaging, it is important in the 3–5 μm band.

A thermal image arises from temperature variations or differences in emissivity within a scene. The thermal contrast is one of the important parameters for IR imaging devices. It is the ratio of the derivative of spectral photon incidence to the spectral photon incidence

$$C = \frac{\partial W / \partial T}{W}.$$

The contrast in a thermal image is small when compared with visible image contrast due to differences in re-flectivity. For a 291 K object in a 290 K scene, it is about 0.039 in the 3–5 μm band and 0.017 in the 8–13 μm band. Thus, while LWIR band may have the higher sensitivity for ambient temperature objects, the MWIR band has the greater contrast.

2.4. Choice of infrared band

In general, the 8–14 μm band is preferred for high performance thermal imaging because of it higher sensitivity to ambient temperature objects and its better transmission through mist and smoke. However, the 3–5 μm band may be more appropriate for hotter object, or if sensitivity is less important than contrast. Also additional differences occur; e.g., the advantage of MWIR band is smaller diameter of the optics required to obtain a certain resolution and that some detectors may operate at higher temperatures (thermoelectric cooling) than it is usual in the LWIR band where cryogenic cooling is required (about 77 K).

Summarising, MWIR and LWIR spectral bands differ substantially with respect to background flux, scene characteristics, temperature contrast, and atmospheric transmission under diverse weather conditions. Factors which favour MWIR applications are: higher contrast, superior clear-weather performance (favourable weather conditions, e.g., in most countries of Asia and Africa), higher

Table 2. Power available in each MWIR and LWIR imaging bands (after Ref. 3).

IR region (μm)	Ground-level solar radiation (W/m^2)	Emission from 290 K blackbody (W/m^2)
3–5	24	4.1
8–13	1.5	127

transmittivity in high humidity, and higher resolution due to ~3 times smaller optical diffraction. Factors which favour LWIR applications are: better performance in fog and dust conditions, winter haze (typical weather conditions, e.g., in West Europe, North USA, Canada), higher immunity to atmospheric turbulence, and reduced sensitivity to solar glints and fire flares. The possibility of achieving higher signal-to-noise (S/N) ratio due to greater radiance levels in LWIR spectral range is not persuasive because the background photon fluxes are higher to the same extent, and also because of readout limitation possibilities. Theoretically, in staring arrays charge can be integrated for full frame time, but because of restrictions in the charge-handling capacity of the readout cells, it is much less compared to the frame time, especially for LWIR detectors for which background photon flux exceeds the useful signals by orders of magnitude.

2.5. Detectors

The figure of merit used for detectors is detectivity. It has been found in many instances that this parameter varies inversely with the square root of both the detector's sensitive area, A, and the electrical bandwidth, Δf. In order to simplify the comparison of different detectors, the following definition has been introduced [5]

$$D^* = \frac{(A\Delta f)^{1/2}}{\Phi_e}(SNR) \qquad (3)$$

where Φ_e is the spectral radiant incident power. D^* is defined as the rms signal-to-noise ratio (SNR) in a 1 Hz bandwidth per unit rms incident radiation power per square root of detector area. D^* is expressed in $cmHz^{1/2}W^{-1}$, which is recently called "Jones". Spectral detectivity curves for a number of commercially available IR detectors are shown in Fig. 3. Interest has centred mainly on the wavelengths of the two atmospheric windows 3–5 μm and 8–14 μm, though in recent years there has been increasing interest in longer wavelengths stimulated by space applications.

Progress in infrared IR detector technology is connected mainly to semiconductor IR detectors, which are included in the class of photon detectors. In the class of photon detectors the radiation is absorbed within the material by interaction with electrons. The observed electrical output signal results from the changed electronic energy distribution. The photon detectors show a selective wavelength dependence of the response per unit incident radiation power. They exhibit both perfect signal-to-noise performance and a very fast response. But to achieve this, the photon detectors require cryogenic cooling. Cooling requirements are the main obstacle to the more widespread use of IR systems based on semiconductor photodetectors making them bulky, heavy, expensive and in-
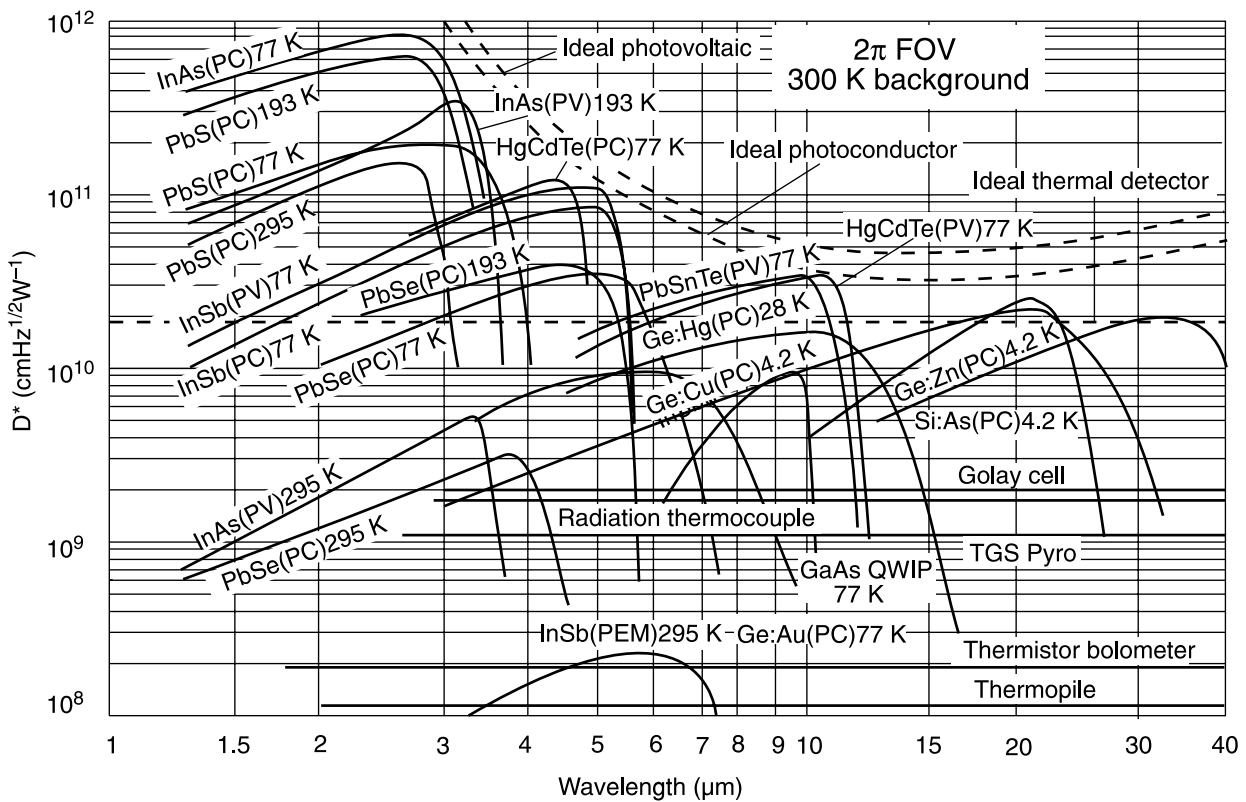

Fig. 3. Comparison of the D^* of various commercially available infrared detectors when operated at the indicated temperature. Chopping frequency is 1000 Hz for all detectors except the thermopile (10 Hz), thermocouple (10 Hz), thermistor bolometer (10 Hz), Golay cell (10 Hz) and pyroelectric detector (10 Hz). Each detector is assumed to view a hemispherical surround at a temperature of 300 K. Theoretical curves for the background-limited D^* for ideal photovoltaic and photoconductive detectors and thermal detectors are also shown (after Ref 5).

convenient to use. Depending on the nature of interaction, the class of photon detectors is further sub-divided into different types. The most important are: intrinsic detectors (HgCdTe, InGaAs, InSb, PbS, PbSe), extrinsic detectors (Si:As, Si:Ga), photoemissive (metal silicide Schottky barriers) detectors, and quantum well detectors (GaAs/AlGaAs QWIPs).

The second class of infrared detectors is composed of thermal detectors. In a thermal detector, the incident radiation is absorbed to change the temperature of material, and the resultant change in some physical properties is used to generate an electrical output. The detector element is suspended on lags, which are connected to the heat sink. Thermal effects are generally wavelength independent; the signal depends upon the radiant power (or its rate of change) but not upon its spectral content. In pyroelectric detectors, a change in the internal spontaneous polarisation is measured, whereas in the case of bolometers a change in the electrical resistance is measured. In contrast to photon detectors, thermal detectors typically operate at room temperature. They are usually characterised by modest sensitivity and slow response but they are cheap and easy to use. Bolometers, pyroelectric detectors, and thermopiles have found the greatest utility in infrared technology. Typical values of detectivities of thermal detectors at 10 Hz change in the range between 10^8 to 10^9 cmHz$^{1/2}$W^{-1}.

Up till the nineties of the 20th century, thermal detectors have been considerably less exploited in commercial and military systems in comparison with photon detectors. The reason for this disparity is that thermal detectors are popularly believed to be rather slow and insensitive in comparison with photon detectors. As a result, the worldwide effort to develop thermal detectors was extremely small relative to that of photon detector. In the last decade, however, it has been shown that extremely good imagery can be obtained from large thermal detector arrays operating uncooled at TV frame rates. The speed of thermal detectors is quite adequate for non-scanned imagers with two-dimensional (2D) detectors. The moderate sensitivity of thermal detectors can be compensated by a large number of elements in 2D electronically scanned arrays. With large arrays of thermal detectors the best values of NEDT, below 0.1 K, could be reached because effective noise bandwidths less than 100 Hz can be achieved.

2.6. Cooling

The signal output of a photon detector is so small that at ordinary temperatures it is swamped by the thermal noise due to random generation and recombination of carriers in the semiconductor. In order to reduce the thermal generation of carriers and minimise noise, photon detectors must be cooled and must therefore be encapsulated. The method of cooling varies according to the operating temperature and the system's logistical requirements. Most 8–14-μm detectors operate at about 77 K and can be cooled by liquid nitrogen. In the field, however, it is more convenient to use

compressed air and a Joule-Thompson minicooler [6]. The operation of Joule-Thompson cooler is based on the fact that as the high-pressure gas expands on leaving a throttle valve, it cools and liquefies. The gas used must be purified to remove water vapour and carbon dioxide which could freeze and block the throttle valve. Specially designed Joule-Thompson coolers using argon are suitable for ultra-fast cool-down.
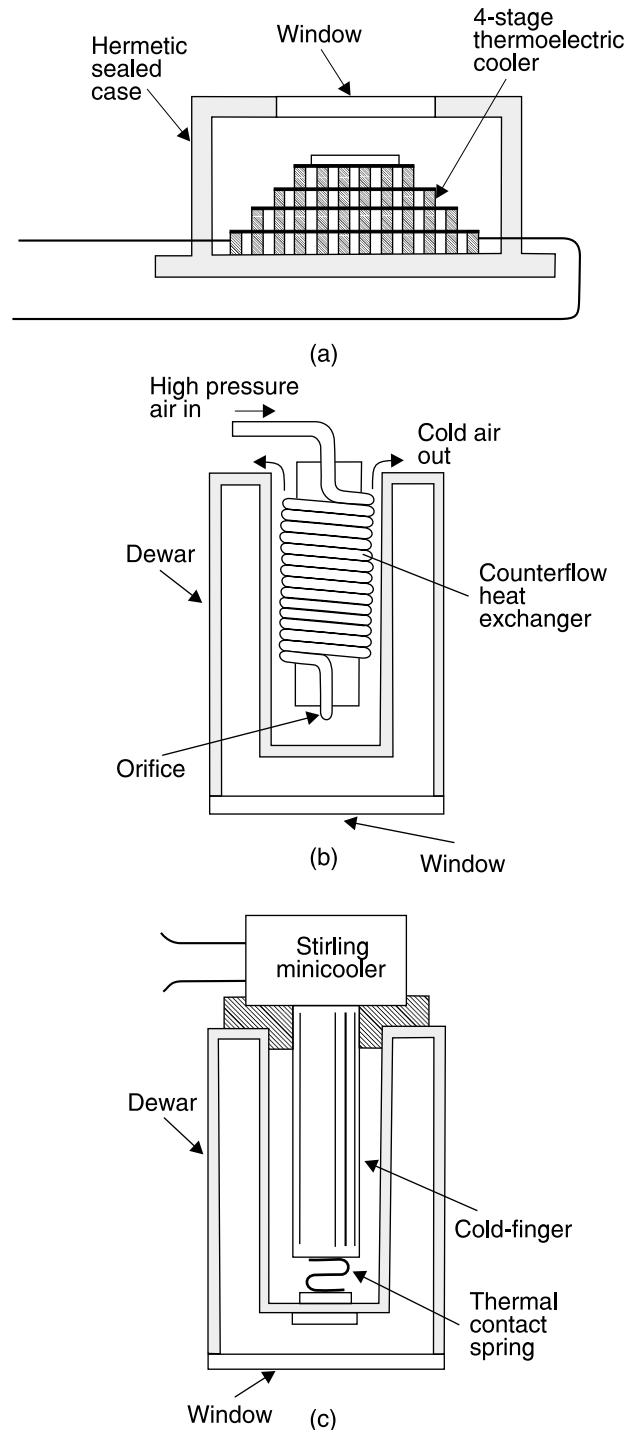

Fig. 4. Three ways of cooling IR detectors: (a) four-stage thermoelectric cooler (Peltier effect), (b) Joule-Thompson cooler, and (c) Stirling-cycle engine.

The use of cooling engines, in particular those employing the Stirling cycle, has increased recently due to their efficiency, reliability and cost reduction. Stirling engine requires several minutes cool-down time; the working fluid is helium. Both Joule-Thompson and engine-cooled detectors are housed in precision-bore dewars into which the cooling device is inserted (see Fig. 4). Mounted in the vacuum space at the end of the inner wall of the dewar, and surrounded by a cooled radiation shield compatible with the convergence angle of the optical system, the detector looks out through an IR window. In some dewars, the electrical leads to detector elements are embedded in the inner wall of the dewar to protect them from damage due to vibration.

Many detectors in the 3–5-μm waveband are thermo-electrically cooled. In this case, detectors are usually mounted in a hermetic encapsulation with a base designed to make good contact with a heatsink.

2.7. Infrared optics

The optical block in IR system creates an image of observed objects in plane of the detector (detectors). In the case of scanning imager, the optical scanning system creates image with the number of pixels much greater that number of elements of the detector. In addition, the optical elements like windows, domes and filters can be used to protect system from environment or to modify detector spectral response.

There is no essential difference in design rules of optical objectives for visible and IR ranges. Designer of IR optics is only more limited because there is significantly fewer materials suitable for IR optical elements, in comparison with those for visible range, particularly for wavelengths over 2.5 μm.

There are two types of IR optical elements: reflective elements and refractive elements. As the names suggest, the role of reflective elements is to reflect incident radiation and the role of refractive elements is to refract and transmit incident radiation.

Mirrors used extensively inside IR systems (especially in scanners) are most often met as reflective elements that serve manifold functions in IR systems. Elsewhere they need a protective coating to prevent them from tarnishing. Spherical or aspherical mirrors are employed as imaging elements. Flat mirrors are widely used to fold optical path, and as reflective prism are often used in scanning systems.

Four materials are most often used for mirrors fabrication: optical crown glass, low-expansion borosilicate glass (LEBG), synthetic fused silica, and Zerodur. Optical crown glass is typically applied in non-imaging systems. It has a relatively high thermal expansion coefficient and is employed when thermal stability is not a critical factor. LEBG, known by the Corning brand name Pyrex, is well suited for high quality front-surface mirrors designed for low optical deformation under thermal shock. Synthetic fused silica has a very low thermal expansion coefficient.

Metallic coatings are typically used as reflective coatings of IR mirrors. There is four types of most often used metallic coatings: bare aluminium, protected aluminium, silver, and gold. They offer high reflectivity, over about 95%, in 3–15-μm spectral range. Bare aluminium has a very high reflectance value but oxidises over time. Protected aluminium is bare aluminium coating with a dielectric overcoat that arrests the oxidation process. Silver offers better reflectance in near IR than aluminium and high
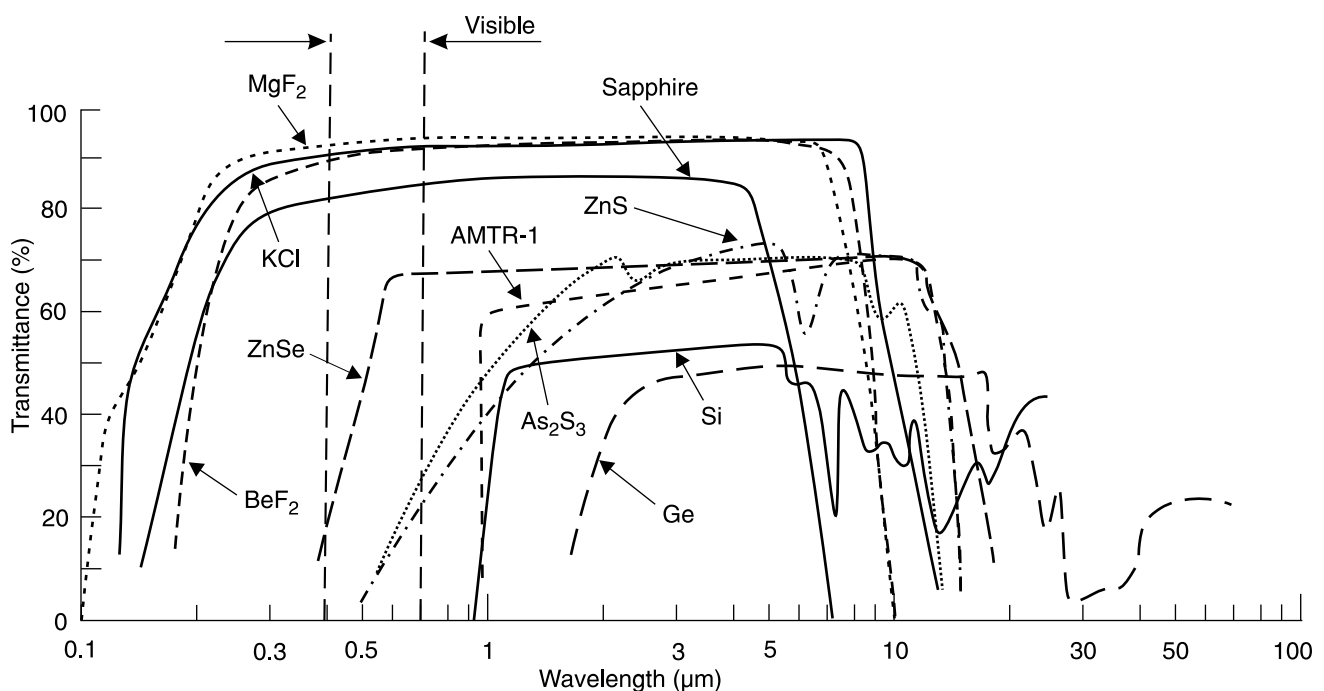

Fig. 5. Transmission range of infrared materials (after Ref. 7).

reflectance across a broad spectrum. Gold is widely used material and offers consistently very high reflectance (about 99%) in the 0.8–50-μm range. However, gold is soft (it cannot be touched to remove dust) and is most often used in laboratory.

The most popular materials used in manufacturing refractive optics of IR systems are: germanium (Ge), silicon (Si), fused silica (SiO₂), glass BK-7, zinc selenide (ZnSe) and zinc sulfide (ZnS). The IR-transmitting materials potentially available for use as windows and lenses are gathered in Table 3 and their IR transmission is shown in Fig. 5.

Germanium is a silvery metallic-appearing solid of very high refractive index (≈ 4), that enables designing of high-resolution optical systems using minimal number of germanium lenses. Its useful transmission range is from 2 μm to about 15 μm. It is quite brittle and difficult to cut but accept a very good polish. Additionally, due to its very high refractive index, antireflection coatings are essential for any germanium transmitting optical system. Germanium has a low dispersion and is unlikely to need colour correcting except in the highest-resolution systems. In spite of high material price and cost of antireflection coatings, germanium lenses are particularly useful for 8–12-μm band. Significant disadvantage of germanium is serious dependence of its refractive index on temperature, so germanium telescopes and lenses may need to be athermalised.

Physical and chemical properties of silicon are very similar to properties of germanium. It has high refractive index (≈ 3.45), is brittle, does not cleave, takes an excellent polish, and has large *dn/dT*. Similarly to germanium, sili-con optics must have antireflection coatings. Silicon offers two transmission ranges: 1–7 μm and 25–300 μm. Only the first one is used in typical IR systems. The material is significantly cheaper than germanium. It is used mostly for IR systems operating in 3–5-μm band.

Single crystal material has generally higher transmission than polycrystalline one. Optical-grade germanium used for the highest optical transmission is n–type doped to receive a conductivity of 5–14 Ω cm. Silicon is used in its intrinsic state. At elevated temperatures semiconducting materials become opaque. As a result, germanium is of little used above 100°C. In the 8–14 μm region, semi-insulating GaAs may be used at the temperatures up to 200°C.

Ordinary glass does not transmit radiation beyond 2.5 μm in IR region. Fused silica is characterised by very low thermal expansion coefficient that makes optical systems particularly useful in changing environmental conditions. It offers transmission range from about 0.3 μm to 3 μm. Because of low reflection losses due to low refractive index (≈ 1.45), antireflection coatings are not needed. It is more expensive than BK-7, but still significantly cheaper than Ge, ZnS and ZnSe, and is a popular material for lenses of IR systems with bands located below 3 μm. BK-7 glass characteristics are similar to fused silica; the difference is only a bit shorter transmission band up to 2.5 μm.

ZnSe is expensive material comparable to germanium; has a transmission range from 2 to 22 μm, and a refractive index about 2.4. It is partially translucent in visible and reddish in colour. Due to relatively high refractive index, antireflection coatings are necessary. The chemical resistance of the material is excellent and superior to germanium and silicon.

Table 3. Principal characteristics of some infrared materials (after Ref. 7).

Material	Waveband (μm)	$n_{4\,\mu m}$, $n_{10\,\mu m}$	dn/dT (10^{-6} K^{-1})	Density (g/cm³)	Other characteristics
Ge	3–5, 8–12	4.025, 4.004	424 (4 μm) 404 (10 μm)	5.33	Brittle, semiconductor, can be diamond-turned, visibly opaque, hard
Si	3–5	3.425	159 (5 μm)	2.33	Brittle, semiconductor, diamond-turned with difficulty, visibly opaque, hard
GaAs	3–5, 8–12	3.304, 3.274	150	5.32	Brittle, semiconductor, visibly opaque, hard
ZnS	3–5, 8–12	2.252, 2.200	43 (4 μm) 41 (10 μm)	4.09	Yellowish, moderate hardness and strength, can be diamond-turned, scatters short wavelengths
ZnSe	3–5, 8–12	2.433, 2.406	63 (4 μm) 60 (10 μm)	5.26	Yellow-orange, relatively soft and weak, can be diamond-turned, very low internal absorption and scatter
CaF₂	3–5	1.410	–8.1 (3.39 μm)	3.18	Visibly clear, can be diamond-turned, mildly hygroscopic
Sapphire	3–5	1.677 (nₒ) 1.667 (nₑ)	6 (o) 12 (e)	3.99	Very hard, difficult to polish due to crystal boundaries
AMTIR-1	3–5, 8–12	2.513, 2.497	72 (10 μm)	4.41	Amorphous IR glass, can be slumped to near-net shape
BF7 (Glass)	0.35–2.3		3.4	2.51	Typical optical glass

ZnS has excellent transmission in the range from 2 µm to 12 µm. It is a high quality crystal that shows only a light yellow colour. Because of relatively high refractive index of 2.25, antireflection coatings are needed to minimise flux reflection. The hardness and fracture strength are very good. ZnS is brittle, can operate at elevated temperatures and also can be used to colour correct high-performance germanium optics.

The alkali halides have excellent IR transmission, however, they are either soft or brittle and many of them are attacked by moisture, making them generally unsuitable for industrial applications.

For more detailed discussion of the IR materials see Harris [8] and Smith [9].

2.8. Night-vision system concepts

Night-vision systems can be divided into two categories: those depending upon the reception and processing of radiation reflected by an object and those which operate with radiation internally generated by an object. The latter systems are described in section 2.9.

The human visual perception system is optimised to operate in daytime illumination conditions. The visual spectrum extends from about 420 nm to 700 nm and the region of greatest sensitivity is near the peak wavelength of sunlight at around 550 nm. However, at night fewer visible light photons are available and only large, high contrast objects are visible. It appears that the photon rate in the region from 800 to 900 nm is five to seven times greater than in visible region around 500 nm. Moreover, the reflectivity of various materials (e.g. green vegetation, because of its chlorophyll content) is higher between 800 nm and 900 nm than at 500 nm. It means that at night more light is available in the NIR than in visual region and that against certain backgrounds more contrast is available.

A considerable improvement in night vision capability can be achieved with night viewing equipment which consists of an objective lens, image intensifier and eyepiece (see Fig. 6). Improved visibility is obtained by gathering more light from the scene with an objective lens than the unaided eye; by use of a photocathode that has higher photosensitivity and broader spectral response than the eye; and by amplification of photo events for visual sensation.
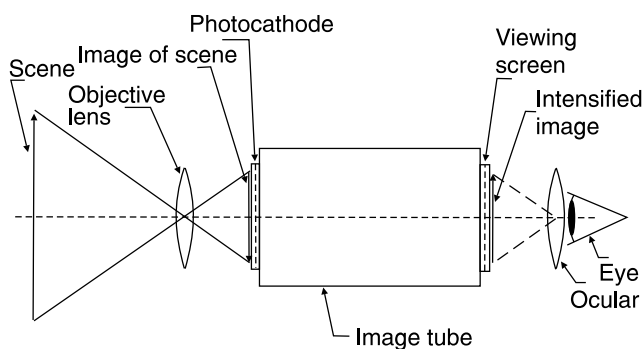

Fig. 6. Diagram of an image intensifier.

2.9. Thermal imaging system concepts

Thermal imaging is a technique for converting a scene's thermal radiation pattern (invisible to the human eye) into a visible image. Its usefulness is due to the following aspects [3]:

- it is totally passive technique and allows day and night operation,
- it is ideal for detection of hot or cold spots, or areas of different emissivities, within a scene,
- thermal radiation can penetrate smoke and mist more readily than visible radiation,
- it is a real-time, remote sensing technique.

The thermal image is a pictorial representation of temperature difference. Displayed on a scanned raster, the image resembles a television picture of the scene and can be computer processed to colour-code temperature ranges. Originally developed (in the sixties last century) to extend the scope of night vision systems, thermal imagers at first provided an alternative to image intensifiers. As the technology has matured, its range of application has expanded and now extends into the fields that have little or nothing to do with night vision (e.g. stress analysis, medical diagnostics). In most present-day thermal imagers, an optically focused image is scanned (mechanically or electronically) across detectors [many elements or two-dimensional (2-D) array] the output of which is converted into a visual image. The optics, mode of scanning, and signal processing electronics are closely interrelated. The number of picture points in the scene is governed by the nature of the detector (its performance) or the size of the detector array. The effective number of picture points or resolution elements in the scene may be increased by an optomechanical scanning device which images different parts of the scene on to the detector sequentially in time.

The performance of a thermal imager is usually specified in terms of temperature resolution. It can be shown, that the temperature sensitivity of an imager, so called noise equivalent temperature difference (NETD), can be given by [10]

$$NETD = \frac{4 f_{\#}^2 (\Delta f)^{1/2}}{A^{1/2} t_{op} M^*},\qquad(4)$$

where $f_{\#}$ is the f-number of the detector optics ($f_{\#} = f/D$, f is the focal length and D is the diameter of the lens), t_{op} is the transmission of the optics, and M^* is the figure of merit which includes not only the detector performance D^* but also the spectral dependence of the emitted radiation, $(\partial S/\partial T)_\lambda$, and the atmospheric transmission t_{at}. It is given by the following equation

$$M^* = \int_0^\infty \left(\frac{\partial S}{\partial T}\right)_\lambda t_{at\lambda} D_\lambda^* d\lambda.\qquad(5)$$

NETD is the difference of temperature of the object required to produce an electric signal equal to the rms noise

at the input of the display [5]. Temperature resolution depends on efficiency of the optical system, responsivity and noise of the detector, and SNR of the signal processing circuitry.

For high sensitivity, the NETD must be low. The sensitivity increases inversely as the square root of the electrical bandwidth. For a given size of IR scene, the electronic bandwidth is inversely proportional to the number of parallel detector elements, and so the thermal sensitivity increases as the square root of the total number of detector elements, irrespective of parallel or serial content of the array.

3. Infrared systems

This section shortly concentrates on selected IR systems and is arranged in order to increase complexity; from smart weapon seekers to space-based systems. A comprehensive compendium devoted to infrared systems was copublished in 1993 by Infrared Information Analysis Centre (IRIA) and the International Society for Optical Engineering (SPIE) as *The Infrared and Electro-Optical Systems Handbook* (executive editors: Joseph S. Accetta and David L. Shumaker).

3.1. Smart weapon seekers

The seeker is the primary homing instrument for smart weapons that include missiles, bombs, artillery projectiles, and standoff cruise missiles. They can be categorised into there groups: passive non-imaging seekers, passive imaging seekers, and active laser guided seekers.

Passive non-imaging seekers use circular optical plate with adjacent transparent and non-transparent parts called reticle that is fixed at the image plane of the imaging optics of the missile head (Fig. 7). A single IR detector, of the size a bit larger than the reticle, is placed just behind it. Location of the point image of the target on the reticle plate changes, even when the target does not change its position, due to rotation of the reticle or rotation of the imaging optics. Therefore radiation emitted by the target generates electrical pulses at the detector output. Pulse duration and phase of these pulses give information about angular position of the target (Fig. 8).

The grandfather of passive IR seekers is the Sidewinder seeker developed in the 1950's; employed vacuum tubes and lead salt single-element detector. During next decades it has been found, that despite their simplicity, the passive non-imaging seekers are very effective for guiding the missiles when the target is on a uniform background. Therefore at present, majority of currently used short-range smart missiles use this type of seekers. However, effectiveness of passive non-imaging seekers decreases significantly for targets on non-uniform background like typical ground military targets or in presence of countermeasures. Therefore the trend of future systems is toward passive imaging seekers.
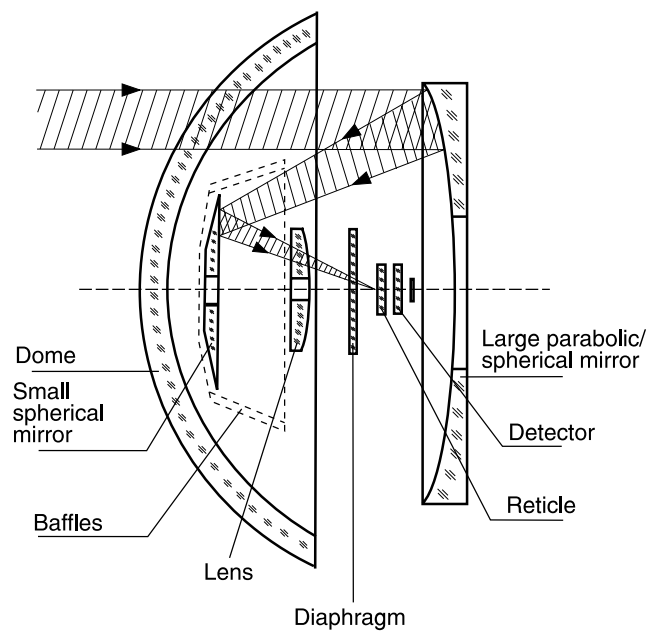

Fig. 7. Optical diagram of a typical passive non-imaging seeker.

Passive imaging seekers have a TV camera or a thermal camera in their optoelectronic head. Location of a target is determined from analysis of the image generated by a camera. Some of the air-to-ground missiles attack and destroy ground targets, particularly large non-movable targets like bridges, bunkers, buildings, etc. However, significant technical limitations exist. First, the seekers using TV cameras can operate only in the day light conditions. Second, it is very difficult to design a thermal camera for the high-speed missiles. Such a camera must be of small-size, very fast operating, reliable, ready to withstand harsh environmental requirements, and of low manufacturing cost. Therefore the imaging missiles using thermal cameras in their optical head are still at a development stage.

Active laser guided seekers can be divided into two subclasses: seekers homing on the irradiated target and seekers irradiated with a laser beam (see Fig. 9). Seekers homing on the irradiated target cooperate with a laser illuminator and use the laser radiation reflected by the target.
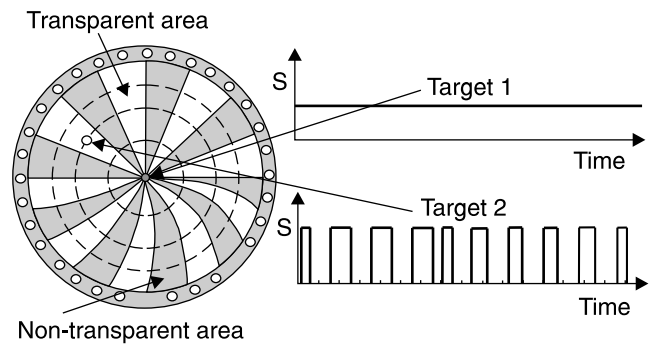

Fig. 8. Exemplary reticle and the signal generated at detector output by a few targets of different location.
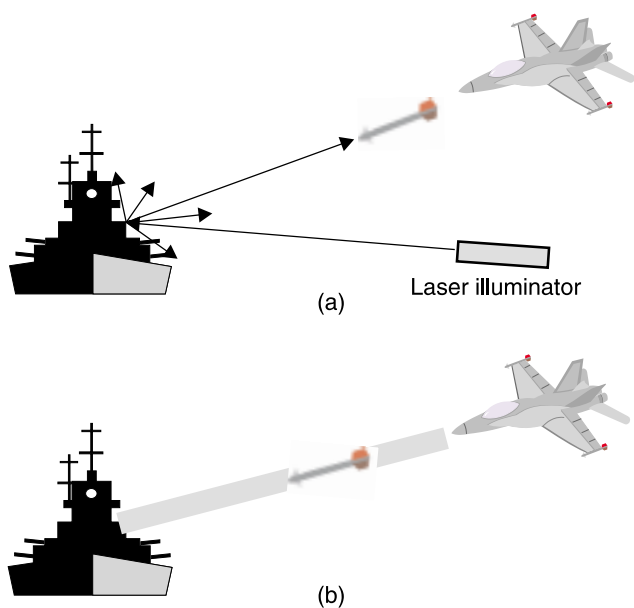
Fig. 9. Principle of work of active laser guided seekers: (a) seekers homing on the irradiated target, and (b) seekers irradiated with a laser beam.

Fig. 10. Representative imaging (staring) seeker architecture (after Ref. 6).

These seekers enable very accurate location of small targets in a highly non-uniform background and are particularly well suited for air-to-ground missiles or bombs. However, it is an active method and employing warning systems or other countermeasures can significantly reduce its effectiveness.

Seekers irradiated with a laser beam are kept on their flight to the target within the beam emitted by the laser illuminator that irradiates the target. Laser radiation, that gives information on target location, comes directly from the illuminator to the sensors at the back of the missile, not after the reflection by the target as in the previous method. Therefore low power illuminators can be used here and the effectiveness of the warning systems is reduced.

The new generation of standoff weapons relies on real-time target recognition, discrimination, tracking, navigation, and night vision. It is predicted that the smart weapons will tend to replace the radar emphasis as stealth platforms are increasingly used for low-intensity conflicts. It is more difficult to perform IR missile warning than radar guided missile warning.

A representative architecture of a staring seeker is shown in Fig. 10. To keep seeker volume, weight, and power requirements low, only the minimum hardware needed to sense the scene is included. We can notice, that the seeker's output is going to a missile-based processor, behind the FPA in the seeker, to perform tracking and aimpoint selection. A concept of seeker's operation includes a standby turn-on, followed by a commit, which cools the FPA. At the beginning, the seeker is locked onto its target by an external sensor or a human. Next, the missile is launched and flies out locked onto its target, matching any target movement. Finally, when the target is close
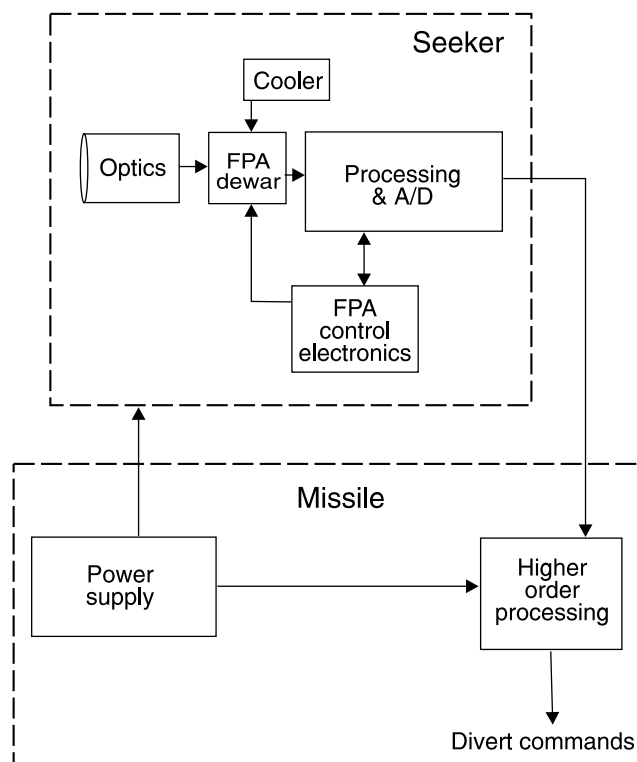
and imaged, the missile chooses an aimpoint and conducts final manoeuvres to get to the target or selects a point and time to fuse and explode.

3.2. Image intensifier systems

The image intensifiers are classed by generation (Gen) numbers. Gen0 refers to the technology of World War II, employing fragile, vacuum-enveloped photon detectors with poor sensitivity and little gain. Gen1 represents the technology of the early Vietnam Era, the 1960s. In this era, the first passive systems, able to amplify ambient starlight, were introduced. Through sensitive, these devices were large and heavy. Gen1 devices used tri-alkali photocathodes to achieve gain of about 1000. By the early 70's, the microchannel-plate (MCP) amplifier was developed comprising more than two million microscopic conducting channels of hollow glass, each of about 10 µm in diameter, fused into a disc-shaped array. Coupling the MCP with multi-alkali photocathodes, capable of emitting more electrons per incident photon, produced GenII. GenII devices boasted amplifications of 20 000 and operational lives of 2500 hrs. Interim improvements in bias voltage and construction methods produced GenII+ version. Substantial improvements in gain and bandwidth in the 1980's heralded the advent of GenIII. Gallium arsenide photocathodes and internal changes in the MCP design resulted in gains ranging from 30 000 to 50 000 and operating lives of 10 000 hrs.

Many candidate technologies could form the basis of a GenIV, ranging from enhanced current designs to com-

pletely different approaches. Among those are devices with a new photocathode that extend spectral response to 1.6 µm and the use of amplifying mechanism other than MCPs. Other potentials include lightweight systems that fuse the outputs from image intensifiers and thermal imagers, and those that couple electron-bombarded CCD arrays – providing sensitivity in the NIR and MWIR regions – with miniature flat-panel displays. The first GanIV tubes demonstrated substantial increase in target detection range and resolution, particularly at extremely low-light levels.

Figure 11 shows the response of a typical GenIII image intensifier superimposed on night sky radiation spectrum. This figure shows also the CIE photopic curve illustrating spectral response of human visual perception system and the GenII response.
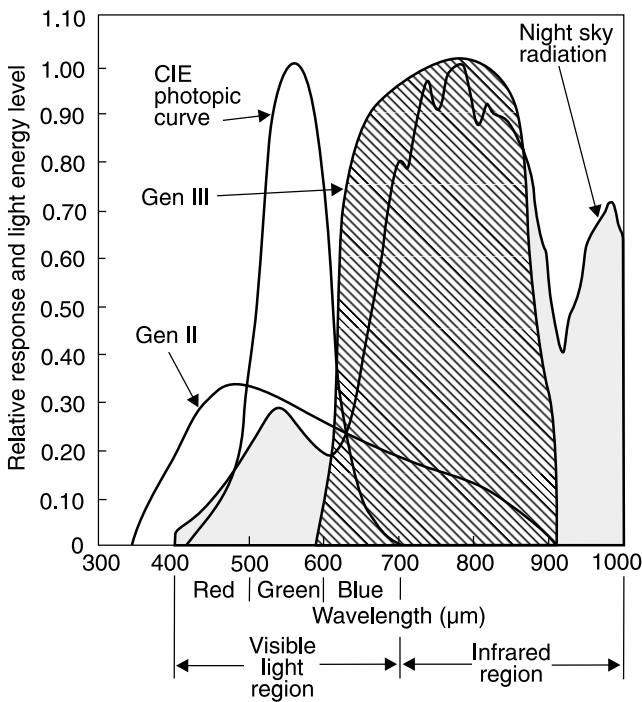

Fig. 11. Image intensifier tube spectral response curves (after Ref.11).

Various implementations of image intensifier tubes have been realised. Phosphor output image intensifiers were reviewed in depth by Csorba [12]. The image is focused onto semitransparent photocathode and photoelectrons are emitted with a spatial intensity distribution, which matches the focused image. In image intensifiers, the electrons are then accelerated towards a phosphor screen where they reproduce the original image with enhanced intensity. Three common forms of image tube are shown in Fig. 12.

In a "proximity-focused" tube, a high electric field (typically 5 kV) and a short distance between the photocathode and the screen, limit spreading of electrons to preserve an image. This form of tube is compact, the image is free from distortion and only simple power supply is required. However, the resolution of such a tube is limited by the field strength at the photocathode and the resolution is highest when the distance between cathode and screen is small.

An electrostatically focused tube is based upon a system of concentric spheres (cathode and anode, typical bias voltage of 15 kV). In practice the electrodes depart radically from the simple spherical concept. Additional electrodes can be introduced to provide focusing control and reduce the image distortion, while fibre optic windows at input and output can be used to improve image quality and provide a better matching to objective and coupling optics. Power suppliers are very simple and lightweight so this type of tube is widely used in portable applications.

A magnetically focused system gives very high-resolution imagers with little or no distortion. The focusing coil, however, is usually heavy and power consuming. For the best picture quality, the power suppliers for both tube and coil must be stable. This type of tube is used in applications where resolution and low distortion are vital and weight and power consumption do not create unacceptable problems.

Image intensifiers are widespread in many military applications. The advent of night vision devices and helmet mounted displays places additional constraints on the helmet, which is now important element of the cockpit dis-
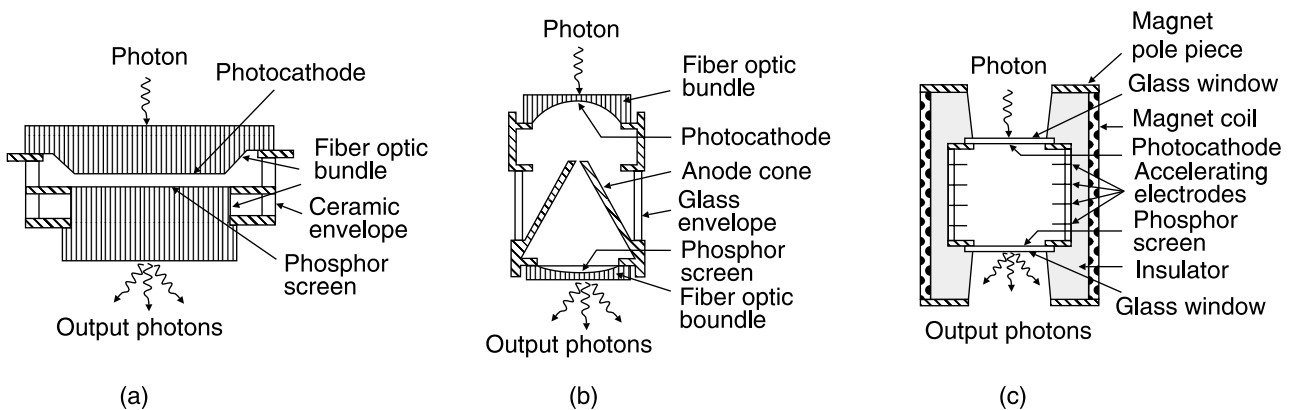

Fig. 12. Cross-sectional diagrams of a variety of image intensifier types: (a) proximity focused, (b) electronically focused, and (c) magnetically focused (after Ref. 12).

Fig. 13. Marconi Avionics' Cats Eyes incorporate an optical combiner assembly for each eye, allowing the pilot to view the cockpit and the outside world directly with the night imagery superimposed on it. Cats Eyes have a 30° field of view and weight 820 g, including the helmet plate (after Ref. 13).

plays system, providing weapon aiming, and other information – such as aircraft attitude and status – to the pilot. For example, Fig. 13 illustrates Marconi Avionics' conventional product produced in large quantities.

3.3. Thermal imaging systems

The basic concept of a modern thermal imager system is to form a real image of the IR scene, detect the variation in the imaged radiation, and, by suitable electronic processing, create a visible representation of this variation analogous to conventional television cameras [14].

Due to existing terminology confusion in literature, we can find at least eleven different terms used as synonyms of the earlier defined thermal imaging systems: thermal imager, thermal camera, thermal imaging camera, FLIR (forward looking infrared), infrared imaging system, thermograph, thermovision, thermal viewer, infrared viewer, infrared imaging radiometer, thermal viewer, thermal data viewer, and thermal video system. The only real difference between the above mentioned terms is that the designations "thermograph", "infrared imaging radiometer", "thermovision", usually refer to thermal cameras used for measurement applications, while other terms refer to thermal cameras used in observation applications. For example, thermographic imagers supply quantitative temperature,

while radiometers provide quantitative radiometric data on the scene (as radiance or irradiance) or process this data to yield information about temperatures.

Thermal imagers have various applications, depending on the platform and user. Most of them are used in military applications. They have often multiple filed of view (FOV) that are user-switchable during operation, what gives both a wide, general surveillance mode as well as a high magnification and narrow field for targeting, designating, or detailed intelligence gathering. Many military thermal imagers are integrated with a TV camera and a laser range finder. TV colour camera is used during daytime conditions due to its superior image quality. Non-military uses include generic search and track, snow rescue, mountain rescue, illegal border, crossing detection, pilot assistance at night or in bad weather, forest fire detection, fire fighting, inspection and discreet surveillance, and evidence gathering. A small but increasing group of thermal imagers enable non-contact temperature measurement and these cameras are used in areas of industry, science, and medicine.

The simplest scanning linear array used in thermal imaging systems, so called focal plane array (FPA), consists of a row of detectors [Fig. 14(a)]. An image is generated by scanning the scene across the strip using, as a rule, a mechanical scanner. At standard video frame rates, at each pixel (detector) a short integration time has been applied and the total charge is accommodated. A staring array is a 2-D array of detector pixels [Fig. 14(b)] scanned electronically.
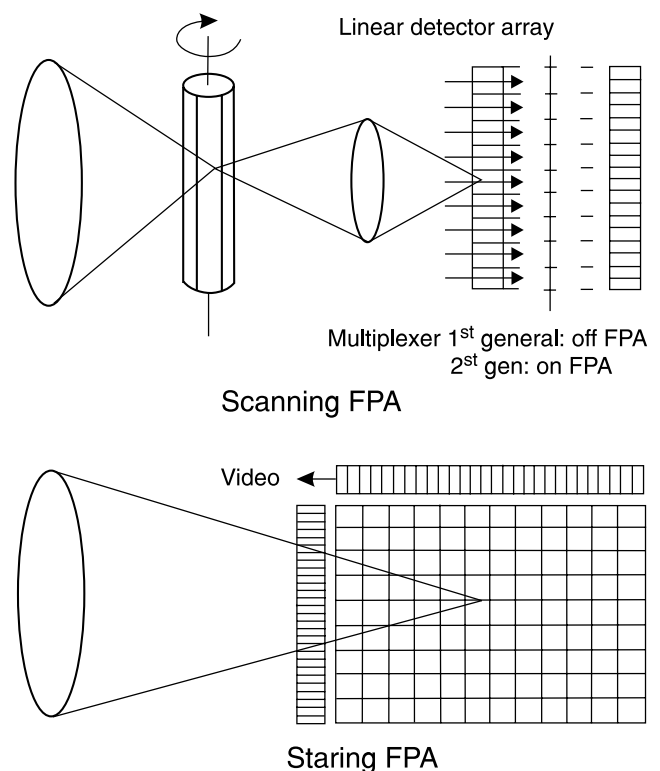

Fig. 14. Scanning (a) and staring (b) focal plane arrays.

The scanning system, which does not include multi-plexing functions in the focal plane, belongs to the first generation systems. A typical example of this kind of detector is a linear photoconductive array (PbS, PbSe, HgCdTe) in which an electrical contact for each element of a multielement array is brought off the cryogenically--cooled focal plane to the outside, where one electronic channel is used at ambient temperature for each detector element. The US common module HgCdTe arrays employ 60, 120, or 180 photoconductive elements depending on the application.

The second generation systems (full-framing systems), being developed at present, have at least three orders of magnitude more elements ($>10^6$) on the focal plane than first generation systems and the detectors elements are configured in a 2-D array. These staring arrays are scanned electronically by circuits integrated with the arrays. These readout integrated circuits (ROICs) include, e.g., pixel deselecting, antiblooming on each pixel, subframe imaging, output preamplifiers, and some other functions. The optics merely focuses the IR image onto the matrix of sensitive elements.

Intermediary systems are also fabricated with multiplexed scanned photodetector linear arrays in use and with, as a rule, time delay and integration (TDI) functions. Typical examples of these systems are HgCdTe multilinear 288×4 arrays fabricated by Sofradir, both for 3–5-μm and 8–10.5-μm bands with signal processing in the focal plane (photocurrent integration, skimming, partitioning, TDI function, output preamplification and some others).

A number of architectures is used in development of IR FPAs [15]. In general, they may be classified as hybrid and
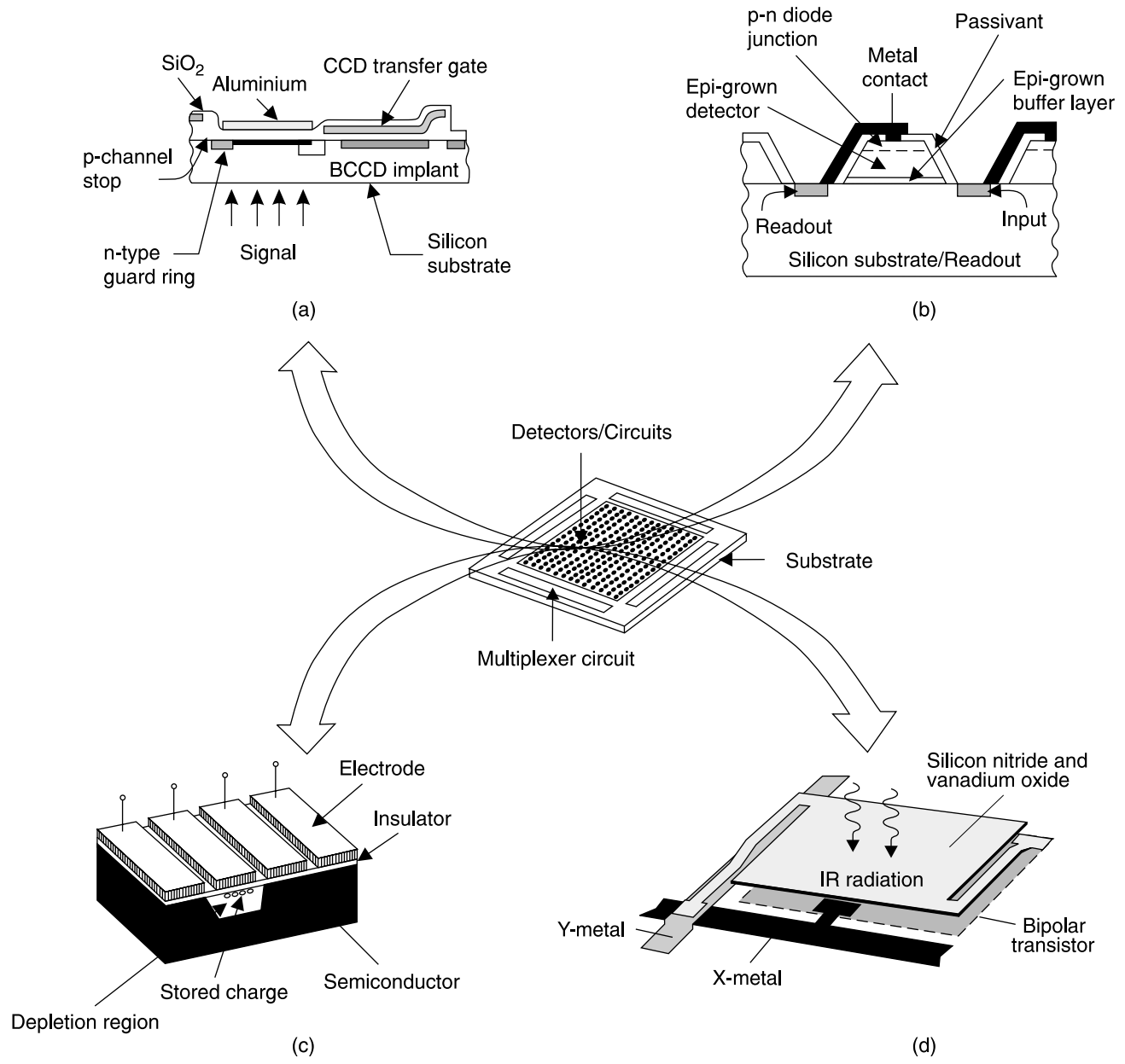

Fig. 15. Monolithic IR FPAs: (a) all-silicon, (b) heteroepitaxy-on-silicon, (c) non-silicon (e.g., HgCdTe CCD), (d) microbolometer.

monolithic ones, but these distinctions are often not as important as proponents and critics state them to be. The central design questions involve performance advantages versus ultimate producibility. Each application may favour a different approach depending on the technical requirements, projected costs and schedule.

In the monolithic approach (Fig. 15), some of the multiplexing is done in the detector material itself than in an external readout circuit. The basic element of a monolithic array is a metal-insulator-semiconductor (MIS) structure. A MIS capacitor detects and integrates the IR-generated photocurrent. Although efforts have been made to develop monolithic FPAs using narrow-gap semiconductors, silicon based FPA technology with Schottky-barrier detectors is the only technology matured to a level of practical use.

Hybrid FPAs detectors and multiplexers are fabricated on different substrates and mated with each other by the flip-chip bonding (Fig. 16) or loophole interconnection. In this case, we can optimise the detector material and multiplexer independently. Other advantages of the hybrid FPAs are near 100% fill factor and increased signal-processing area on the multiplexer chip. In the flip-chip bonding, the detector array is typically connected by pressure contacts via indium bumps to the silicon multiplex pads. The detector array can be illuminated from either the frontside or backside (with photons passing through the transparent detector array substrate). In general, the latter approach is most advantageous. When using opaque materials, substrates must be thinned to 10–20 μm in order to obtain sufficient quantum efficiencies and reduce the crosstalk.

Two types of silicon addressing circuits have been developed: CCDs and complementary metal-oxide-semiconductor (CMOS) switches. In CCD addressing circuits the photogenerated carriers are first integrated in the well formed by a photogate and subsequently transferred to slow (vertical) and fast (horizontal) CCD shift registers [16].

An attractive alternative to the CCD readout is coordinative addressing with CMOS switches. The advantages of CMOS are that existing foundries. Design rules of 0.25 μm are in production with pre-production runs of the 0.18 μm design rules. At present, CMOS with minimum feature ≤0.5 μm is also enabling monolithic visible CMOS imagers.

The minimum resolvable temperature difference (MRTD) is currently considered as the most important parameter of thermal imaging systems (see STANAG No. 4349) [17]. MRTD enables us to estimate probability of detection, recognition, and identification of military targets knowing MRTD of the evaluated thermal imager. Military standards determining testing the thermal imaging systems usually specify that MRTD values for a set of spatial frequencies of the tested imager must be lower than certain values if the imager is to pass the test.
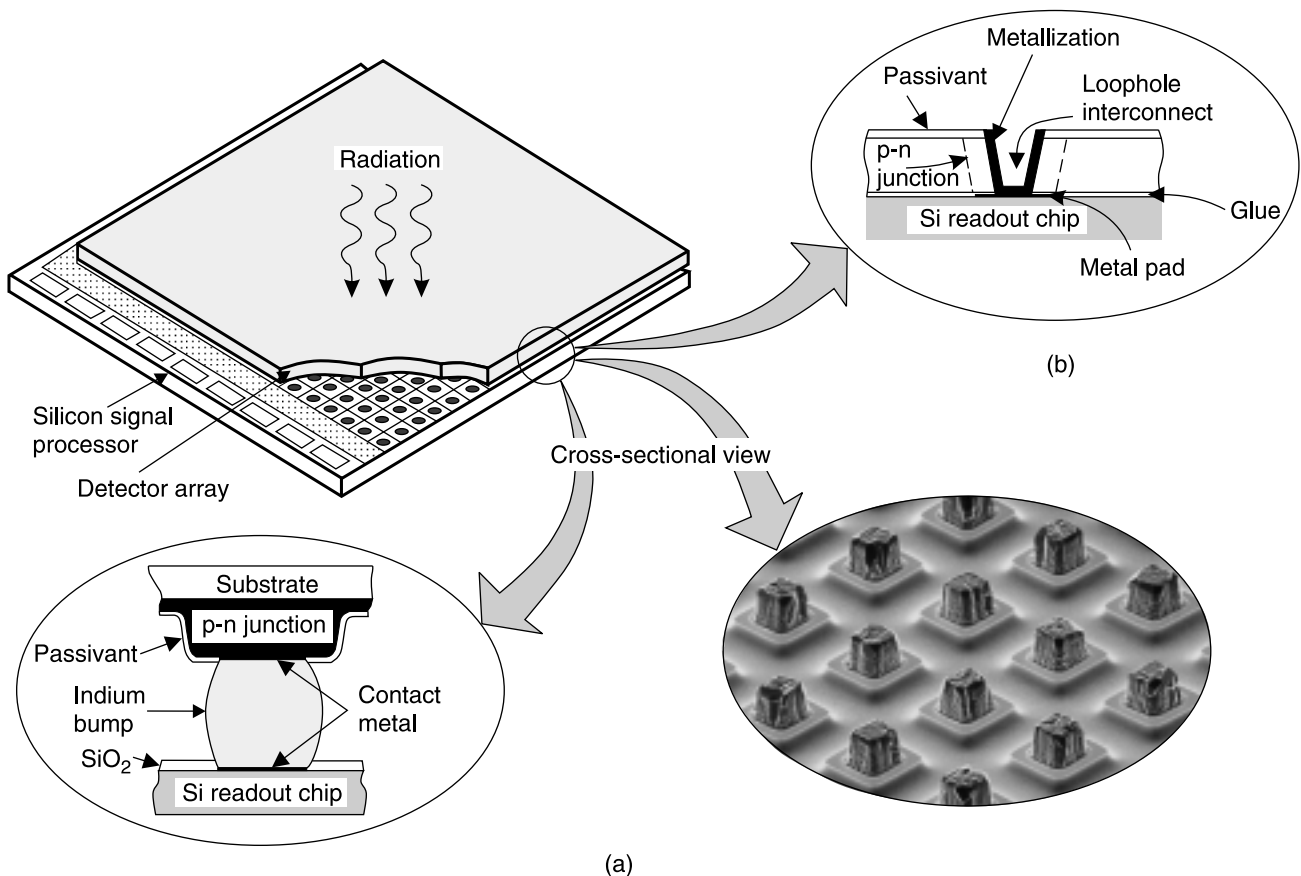

Fig. 16. Hybrid IR FPA interconnect techniques between a detector array and silicon multiplexer: (a) indium bump technique, and (b) loophole technique.
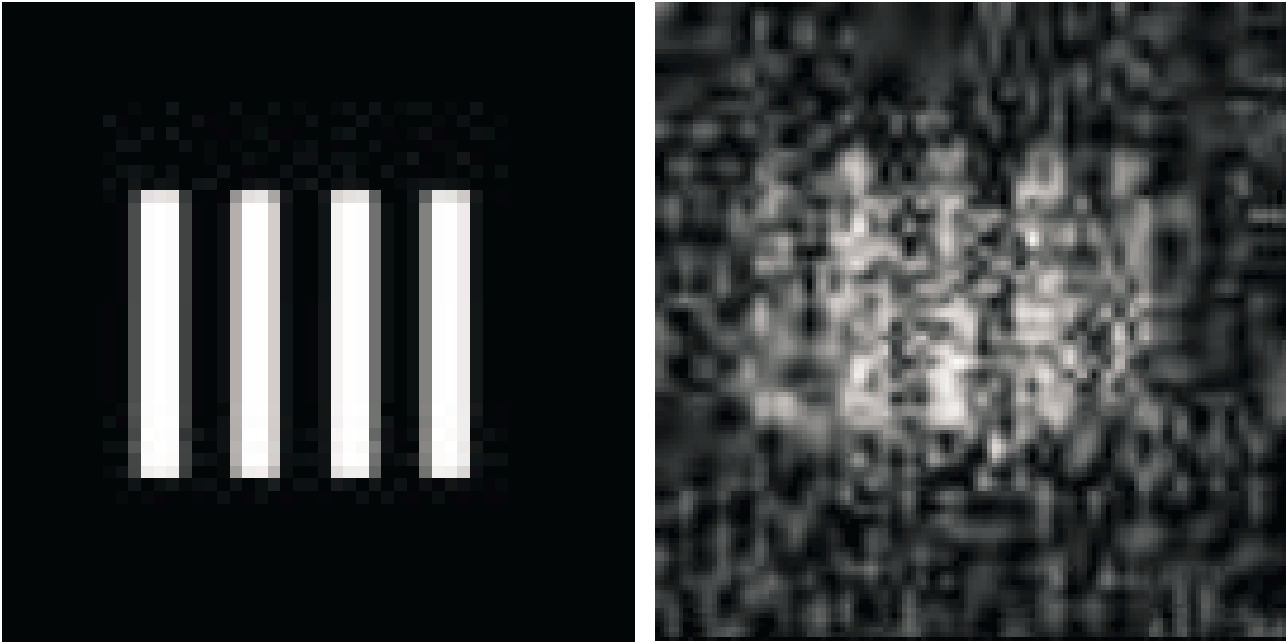
Fig. 17. Image of standard 4-bar target used during MRTD measurement: (a) high temperature difference between target and background, and (b) low temperature difference between target and background.

The MRTD is a subjective parameter that describes ability of the imager-human system for detection of low contrast details of the tested object. It is measured as a minimum temperature difference between the bars of the standard 4-bar target and the background required to resolve the thermal image of the bars by an observer versus spatial frequency of the target (see Fig. 17). The measurement results of typical military thermal imagers for airborne surveillance are shown in Fig. 18.

3.3.1. Infrared cameras vs. FLIR systems

Historically, a "camera" includes neither the storage medium nor the display, while "camera system" includes the complete package. At present, the manufacturers offer an optional recording medium (usually CD ROM), display, and electronics for the display. For example, Fig. 19 is a photograph of Inframetrics microbolometer IR camera ThermaCam 395.

"FLIR" is archaic sixties jargon for forward-looking infrared to distinguish these systems from IR line scanners, which look down rather than forward. Conversely, most sensors that do look forward are not considered to be FLIRs (e.g. cameras and astronomical instruments). The term "FLIR" should be eliminated from IR techno-speak, but is still used and is likely to remain in the jargon for a while.

It is difficult to explain the difference between camera and a FLIR system. In general, FLIRs are designed for specific applications and specific platforms, their optics is integrated into the package, and they are mostly used by people. Cameras usually rely on "imaging" of a "target" and
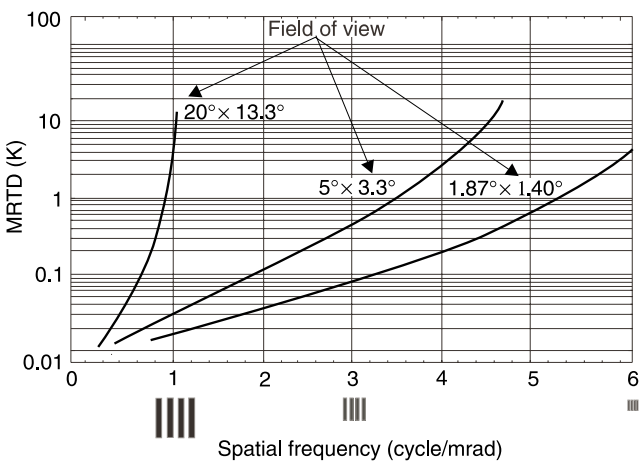

Fig. 18. MRTD of exemplary military thermal camera used in airborne surveillance.

Fig. 19. Inframetrics microbolometer IR camera ThermaCam 395 (photography courtesy of Inframetrics).

they are designed for generic purposes, without much consideration for form and fit; they can be used with many different fore optics and are often used by computers and machines (not just people).

The term "FLIR" usually implies military or paramilitary use, air-based units, and scanners. The FLIR provides automatic search, acquisition, tracking, precision navigation, and weapon delivery functions. A typical FLIR is comprised of four line replaceable units, such as a FLIR optical assembly mounted on a gyro-stabilised platform, electronics module containing all necessary electronics circuits and cryogenically cooled detector array, a power supply unit, and control and processing assembly.

Figure 20 shows representative camera architecture with three distinct hardware pieces: a camera head (which contains optics, including collecting, imaging, zoom, fo-
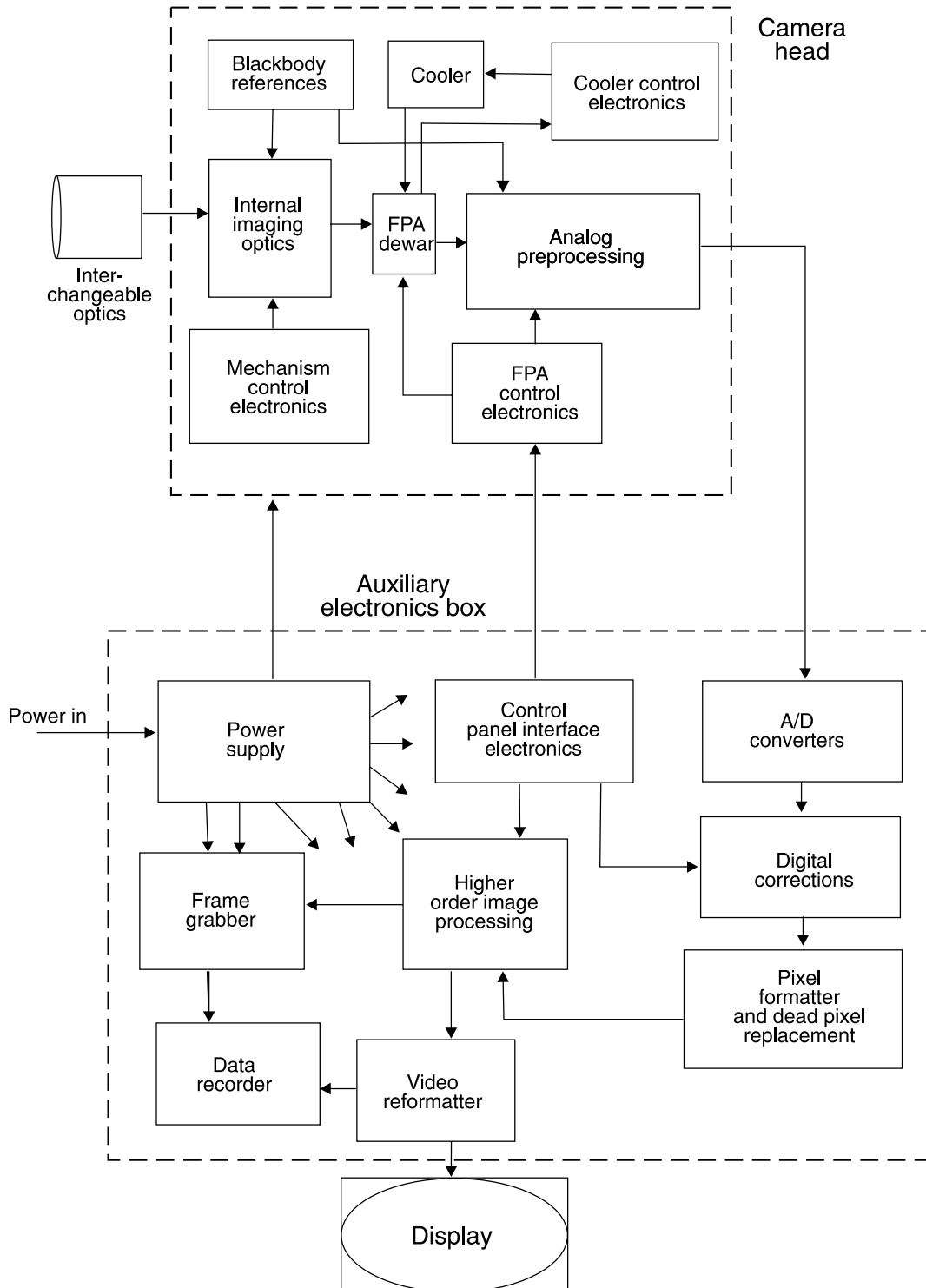

Fig. 20. Representative IR camera system architecture (after Ref. 6).

cusing, and spectral filtering assembles), electronics/control processing box, and the display. Electronics and motors to control and drive moving parts must be included. The control electronics usually consist of communication circuits, bias generators, and clocks. Usually camera's sensor (FPA) needs cooling and therefore some form of cooler is included, along with its closed-loop cooling control electronics. Signal from FPA is of low voltage and amperage and requires analogue preprocessing (including amplification, control, and correction), which is located physically near the FPA and included in the camera head. Often, the A/D is also included here. For user convenience, the camera head often contains the minimum hardware needed to keep volume, weight, and power to a minimum.

Typical costs of cryogenically cooled imagers around $50000 restrict their installation to critical military applications allowing conducting of operations in complete darkness. Moving form cooled to uncooled operation (e.g., using silicon microbolometer or BST pyroelectric arrays) reduces the cost of an imager to below $20000. The cost of pyroelectric vidicons is usually a few thousand dollars, however, they have NETD of 0.5°C (although some of them are now reported as low as 0.2°C) and typically poor image quality compared with full frame staring arrays. They present a major departure from camera architecture presented in Fig. 20.

Cameras usually produce high-quality imagers with NETDs of 0.05 to 0.1°C. Details and resolution vary by optics and focal planes. A good camera produces an image akin to that of a black and wide television.

In the 1960's, the earliest FLIRs were linear scanners. In the 1970's, first-generation common modules (including a dewar containing 60, 120, or 180 discrete elements of photoconductive HgCdTe) were introduced. Next generation of FLIRs employed a dense linear array of photovoltaic HgCdTe, usually 2(4)×480 or 2(4)×960 elements in TDI for each element. At present, these systems are replaced by full-framing FLIRs that employ staring arrays (PtSi, HgCdTe, InSb and QWIP).

FLIRs are usually in several discrete packages referred to as line replaceable units (LRUs) such as: scanner head, power supply, image processor, recorder, display, and controls. They have the form of boxes spread around the host platform. The controls and display must be mounted in the cockpit with the humans. A representative FLIR architecture with the video signal output (to support LRU for image and higher-order processing) is shown in Fig. 21. Many systems depart significantly from the architecture of Fig. 21.

FLIRs usually use telescopes in the sense that the lens system is focused at a distance very much larger than the focal length. Characteristics such as FOV, resolution, element size, and spatial frequency are expressed in angular units. By convention, FOV is expressed in degrees, resolution in milliradians, spatial frequency in cycles per milliradian, and noise in units of temperatures.

Worldwide, there are over 100 different FLIR systems in operation. The most important of them are described in the literature [6,14]. Several FLIRs integrate a laser ranger or target designator.

Figure 22 shows Falcon Eye, a representative conformal FLIR. This IR device receiver unit features a special gimbals' set that allows the 5-inch ball to rotate side-to-side and up and down. This head-steered FLIR adds an excellent degree of tactical flexibility and night situation awareness by allowing the pilot to look in any direction – including directly above the aircraft.

Recent outgrowths of military FLIRs are the infrared search and track (IRST) systems. They are subset or class of passive systems whose objective is to reliably detect, locate, and continuously track IR-emitting objects and targets in the presence of background radiation and other disturbances. They are used in radar-like manner (usually with radar-like display) to detect and track objects. Most of the current research in IRST systems is concentrated in signal processing to extract target tracks from severe clutter.

Another group outgrowths of military thermal imagers are airborne line scanners. These are one-dimensional scanning systems that enable creation of a two dimensional thermal image of the observed scenery only when the system is moving. In contrast to typical thermal imagers with FOV not higher than about 40°, the airborne line scanners can provide field of view up to about 180°. Due to wide FOV airborne thermal scanners are widely used in military aerial reconnaissance.

3.4. Space-based systems

The formation of NASA in 1958, and development of the early planetary exploration program, was primarily responsible for the development of the modern optical remote sensing systems, as we know them today. During the 1960s optical mechanical scanner systems became available that made possible acquisition of image data outside the limited spectral region of the visible and NIR available with film. "Eye in the sky" was the first successfully flown long wavelength sensor launched in 1967. A major milestone was the development of the Landsat Multispectral Scanner since it provided the first multispectral synoptic in digital form. The period followed the launch of Landsat-1 in 1972 stimulated the development of a new series of air- and spaceborne sensors. Since that time, hundreds of space-based sensors have been orbited.

The main advantages of space IR sensors are as follows [6]:
- ability to tune the orbit to cover a ground swath in optimal spatial or temporal way,
- lack of atmospheric effects on observation.
- global coverage,
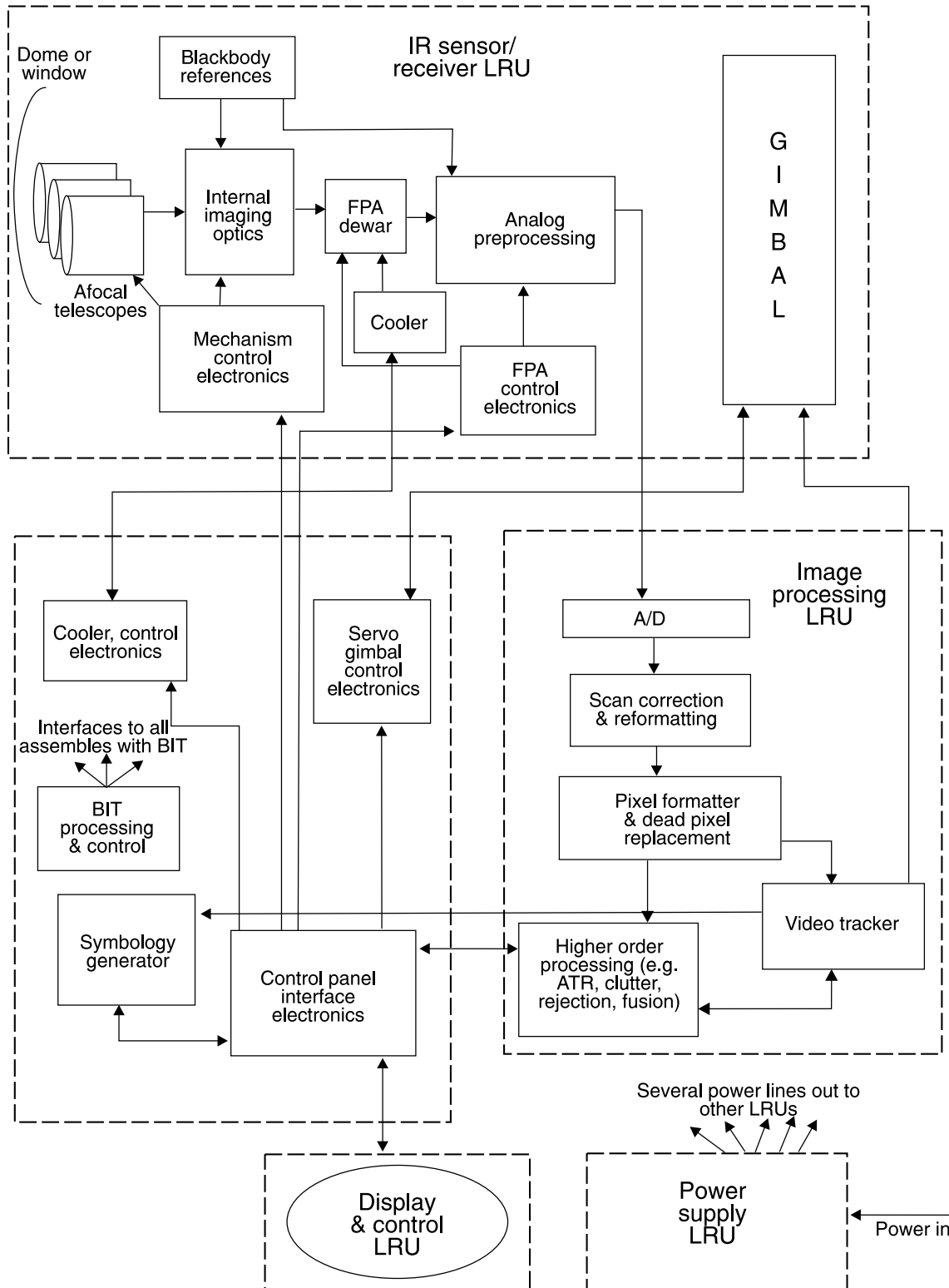- ability to engage in legal clandestine operations.

Fig. 21. Representative FLIR system architecture (after Ref. 6).

Hitherto, anti-satellite weapons do not exist, so satellites are relatively safe form attack. The disadvantages of satellite systems are protracted and excessive costs of fabrication, launch, and maintain satellites. Moreover, such operations as repair and upgrade are difficult, expensive, and usually not possible.

The space-based systems installed on space platforms usually perform one of the following functions [6]: military/intelligence gathering, astronomy, Earth environmental/resources sensing, or weather monitoring. So, these functions can be classified as forms of Earth remote sensing and astronomy.

Fig. 22. Falcon Eye, a representative conformal FLIR (after Ref. 18).

Figure 23 shows representative space sensor architecture. It should be stressed, however, that many individual space sensors have not got this exact architecture.

Intelligence and military services from wealthy nations have long employed space-based sensors to acquire information. A satellite-borne IR warning receiver, designed to
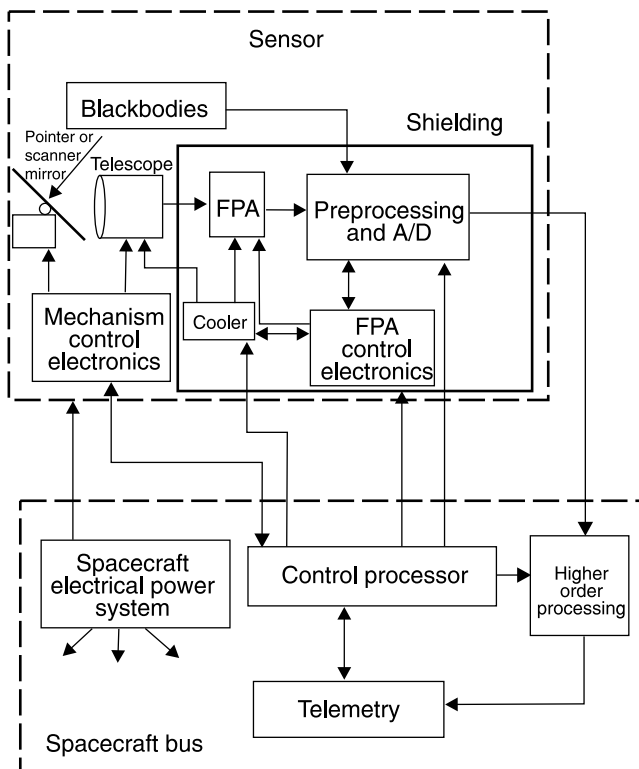

Fig. 23. Representative space sensor architecture (after Ref. 6).

detect intercontinental ballistic missiles is a strategic system that protects a large area, or nation. The U.S. spends about $5 billion per year on space reconnaissance [6]. Although the cold war is over, the long-term strategic monitoring to access military and economic might is still important. Intelligence gathering of crop data and weather trends from space has also been used by hunger relief organisations to more effectively forecast droughts and famines. The military has also space-based surveillance for missile launches and additionally, space basing provides excellent viewing geometries for global events as nuclear explosions and environmental changes that the military is concerned about.

Imaging with IR FPAs provides increasingly detailed and quantitative information about relatively cool objects in space of our galaxy and beyond. Dwarf stars, for example, or giant Jupiter-like planets in other distant solar systems do not emit much visible and ultraviolet light, so they are extremely faint at these wavelengths. Also, the longer IR wavelengths can penetrate dusty and optically opaque nebulous molecular clouds in interstellar space where new stars and planetary systems are forming.

There are several unique reasons for conducting astronomy in space [6]:
• to eliminate the influence of absorption, emission and scattering of IR radiation,
• to answer basic cosmological and astronomical questions (e.g. formation of stars, protoplanetary disks, extra-solar planets, brown dwarfs, dust and interstellar media, protogalaxies, the cosmic distance scale, and ultra-luminous galaxies).
• to observe the Earth's environmental (detecting the subtle changes indicating environmental stresses and trends).

4. Non-contact thermometers

Infrared thermometers always measure temperature indirectly in two stages:
• measurements of radiation power in one or more spectral bands,
• determination of an object temperature on the basis of the measured radiometric signals.

Even simple IR thermometers usually consist of five or more blocks (see Fig. 24). An optical objective is usually used to increase the amount of radiation emitted by the tested object and to limit thermometer field of view. The signal at the output of the detector is typically amplified, converted into more convenient electronic form and finally digitised. A separate visualisation block is typically used for presentation of the measurement results.

The IR thermometers can be divided into a few groups according to different criteria: presence of an additional co-operating source, number of system spectral bands, number of measurement points, width of system spectral bands and transmission media.
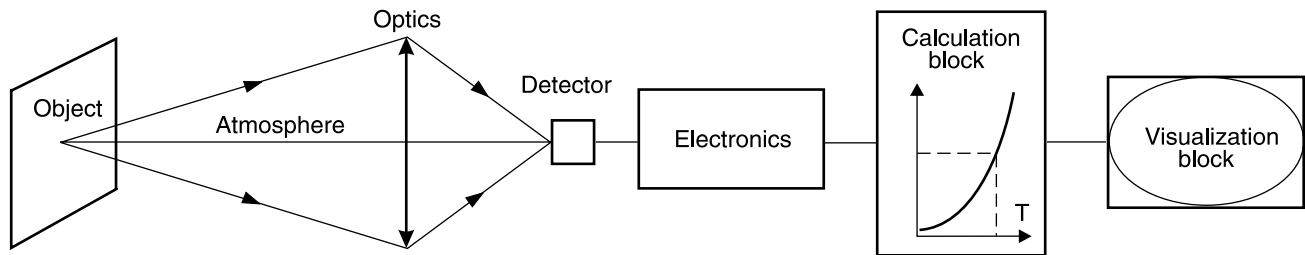
Fig. 24. General diagram of a simple non-contact thermometer (after Ref. 19).

In passive system, the object temperature is measured knowing radiation power emitted by the object in one or more spectral bands. With active system we can get some information about emissive properties of the tested object by using an additional co-operating source that emits radiation directed to the tested object and measuring the reflected radiation. They are active systems. Active thermometers are more sophisticated, more expensive and so far only in few applications they can really offer better accuracy than passive systems. Therefore, nowadays, almost all, non-contact thermometers are passive ones.

In the passive singleband systems the object temperature is determined using system calibration chart derived from radiometric calculation of the output signal as a function of blackbody temperature. The temperature of non-blackbody objects can be corrected if only their emissivity over the spectral band is known.

The ratio of the power emitted by a graybody at two different wavelengths does not depend on the object emissivity but only on the object temperature. In passive dualband systems, the object temperature is usually determined using a calibration chart that represents a ratio of the emitted power in two bands as a function of the object temperature.

At present, at least 95% of systems available commercially on market are passive singleband systems; passive dualband systems are rather rarely used; passive multiband systems are still at a laboratory stage of development.

According to number and location of measurement points, the infrared thermometers can be divided into pyrometers, line scanners and thermal cameras. Pyrometers enable temperature measurement of only a single point or rather of a single sector (usually a circle or a square) of the surface of the tested object. Line scanners enable temperature measurement of many points located along a line. Thermal cameras enable temperature measurement of thousands of points located within a rectangle, square or circle, and create a 2-D image of temperature distribution on this area.

Most commercially available non-contact thermometers are pyrometers (see Fig. 25). They are small, light and low-cost systems that found numerous applications in industry, science, etc. Line scanners are especially suitable for temperature measurement of moving objects and found applications in automotive industry, welding, robotics etc. Thermal cameras offer the greatest capabilities of all discussed types of non-contact thermometers. In spite of their high price, they found numerous applications such as con-

Fig. 25. M90V pyrometer manufactured by Mikron Instrument Company, Inc. (photography courtesy of Mikron Inc.).

trol of electrical supply lines, heat supply lines, civil engineering, environmental protection, non-destructive testing and many others.

The fixed, inflexible configuration of non-contact thermometer (like that presented in Fig. 25) is not a good solution in situations when direct sighting due to obstructions is impossible, significant interference is present and electronics must be placed in safe distance, or very high temperatures exist. In such situations it is better to use flexible fibre thermometers without the optics block.

5. Radiometers

In general, the IR thermometers discussed in the previous section can be treated as a class of radiometers because they determine temperature on the basis of the signal generated by the radiant flux coming to the detector. However, the IR thermometers are designed to measure only temperature and it is usually not possible to use them to measure radiant flux. In our definition, radiometer is an instrument designed to measure quantities of infrared radiation, radiant properties of materials, or infrared detector parameters.

The IR radiometers can be divided into few groups according to different criteria: measured quantity, number of spectral bands, number of measurement points.

Radiometers enable measurement of such quantities of infrared radiation like radiant power, radiant energy, radi-

ant intensity, radiance, irradiance, radiant exposure; radiant properties of materials (like emissivity, reflectance and transmittance); parameters of infrared detectors (like responsivity, and detectivity). However, all these features are possessed only by a small group of radiometers that are generally bulky, laboratory type and expensive systems. Their extremely versatility is usually achieved by modular approach coupled with an extensive selection of accessories and powerful application software packages what enables the user to tailor a turn-key system to their exact requirements as well as insure expandability in the future. On the opposite side there are radiometers of design optimised for measurements of only a single quantity. Optical power meters are the prime example of radiometers from the latter group (see Fig. 26).

Fig. 26. Optical power meter Model 1830 (photography courtesy of Newport Inc.).

According to the criterion of number of spectral bands the IR radiometers can be divided into the following groups: single-band radiometers, dual-band radiometers and multiband radiometers. Multiband radiometers enable measurements of one of the above mentioned radiometric quantities in at least three separate spectral bands. When the spectral bands are narrow and their numbers is high enough, then the multiband radiometers are termed spectroradiometers. In contrast to situation in non-contact thermometry where the singleband IR thermometers dominate on the market, the spectroradiometers have found wide area of applications and the most popular group of the radiometers.

The key component of any spectroradiometer is a module that can be termed spectral band selector. Its task is to select the desired spectral band from the incoming radiation. This task is achieved by the use of three methods: variable filters, monochromators, and Fourier Transform (FT) interferometers.

The transmission wavelength of circular (linear) variable filters (VF) changes continuously (discretely) with position of the fraction of the filter. Simplicity of design is a great advantage of this type of spectroradiometers; it enables design of small size, reliable, high speed and mobile systems. However, using variable filters it is not possible to

achieve very good spectral resolution (typically about 2% of the wavelength). Next, because the system must measure output radiation selected by variable filter of narrow spectral band and low transmission coefficient it is necessary to use cooled infrared detectors. Cooled sandwich InSb/HgCdTe enabling measurement in the spectral range 2.5 μm to 15 μm is a typical option for this type of spectroradiometers.

Monochromator is an optical instrument that uses a dispersing component (a grating or a prism) and transmits to the exit slit (optionally directly to detector) only a selected fraction of the radiation incoming to the entrance slit. The centre wavelength of the transmitted spectral band can be changed within the instrument spectral region by rotation of the dispersing element. Dispersing prisms, or more often gratings are used as the dispersing elements in monochromators.

The Michelson interferometer is the spectral band selector in FT spectroradiometers. The interferometer is usually built as an optical instrument consisting of a beam splitter and two flat mirrors arranged such as to recombine the two separated beams back on the same spot at the beamsplitter. One of the mirrors moves linearly in order to produce variable optical interference.

The Michelson interferometer can also be seen as a modulator. From a constant spectral radiation input, a temporal modulation occurs at the detector having a unique modulation frequency for each wavelength of radiation. The modulation frequency can be scaled via the velocity of the mirror movement. This modulated signal registered by the detector is called the interferogram. It is digitised at the rate of at least two times the maximum modulation frequency and a mathematical operation, the Fourier Transform, is applied to retrieve the spectral distribution of the input radiation. Calibration with a known source is required in order to obtain quantitative radiometric results.

FT spectroradiometers are characterised by a very good spectral resolution and very good sensitivity, better than offered other types of spectroradiometer. Very good spectral resolution is the effect of use of the interferometer as a spectral selector. Very good sensitivity originates from the fact that the detector is irradiated not only by the radiation from a desired narrow spectral band but also by a full spectrum of radiation coming to the interferometer input. This feature enables design of high-speed, high spectral resolution FT spectroradiometers using non-cooled or thermoelectrically cooled detectors (typically HgCdTe detectors) instead of bulky liquid nitrogen cooled detectors needed in the variable filter or dispersive spectroradiometers. However, performance of the FT spectroradiometers can be severely reduced even by a very small non-alignment of the optical system what makes this type of spectroradiometers inherently sensitive to shocks and vibrations. Therefore FT spectroradiometers were for the last few decades considered as rather laboratory type equipment that cannot be used in field applications. However, at present, this opinion is outdated as
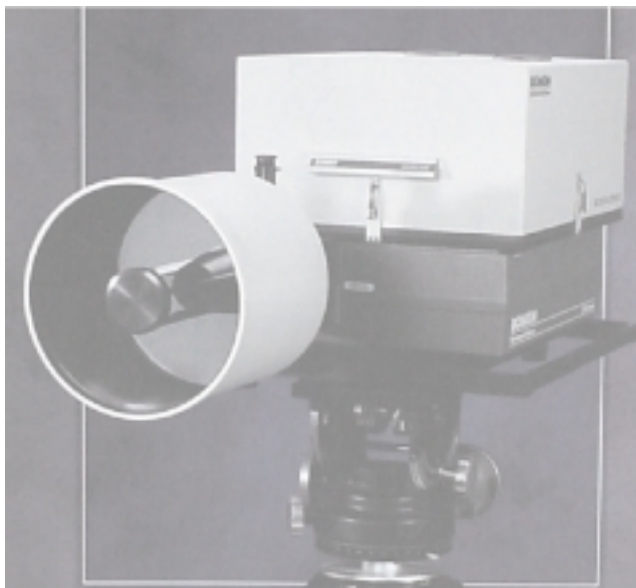
Fig. 27. FT spectroradiometer (photography courtesy of Bomem Inc.).

fully mobile FT spectroradiometers are on the market (see Fig. 27).

Great majority of commercially available spectroradiometers are the systems enabling measurement of the spectral distribution of radiation emitted or reflected by a single spot and these systems can be termed the spot radiometers. There exists also another group termed the imaging spectroradiometers because these systems offer some imaging capabilities.

Optical system of any monochromator creates, at the output plane, a series of adjacent images of the input slit corresponding to wavelength. At one time only, one of these images fits to the exit slit and this radiation is mea-

sured by the detector located behind the stationary exit slit. Rotation of the dispersing element causes movement of these series images and the radiation from each of them can be measured. If we put an array detector at the output plane of the monochromator optical system, instead of traditional configuration (an exit slit and a single detector behind), then this series adjacent images of the input slit corresponding to wavelength will be focused on different parts of the array detector. Therefore by use of an array detector we could expect possibility of simultaneous measurement of radiation spectrum of different spots within the input slit. However, due to significant aberrations (curved output field, astigmatism) of the optical systems this cannot be achieved in standard non-imaging spectroradiometers. The aberrations create situation when only one image from the series – the image that fits the exit slit – is horizontally sharp. By use of modified optical system with corrected curved output field and astigmatism we can get sharp images of the input slit corresponding to wavelength focused on different parts of the array detector what enables simultaneous measurement of spectrum.

The second class of imaging spectroradiometers are multiband (in literature multispectral or hyperspectral) imaging systems generating simultaneously 2-D images of the observed scenery in a number of spectral bands, where this number can vary from a few bands to over a hundred.

High number of spectral bands being simultaneously recorded is typically achieved by use of a number of dichroic beam splitters, gratings, and linear detectors of different spectral sensitivity regions. The beam splitters separate the incoming parallel polychromatic beam into a few beams of separate spectral bands; for example see Table 4. The grating separates the beams further and finally all these spectrally separated beams are focused by output optical objectives at different elements of the linear detectors.

Table 4. Basic parameters of the GER EPS-H imaging spectroradiometers (see www.ger.com/epsh.html) (permission from the GER Inc.).

	Parameters	
Scanner	Kennedy large-size scan head	
Spectrometers	1 Si (VIS/NIR)	76 channels; 0.43–1.05 μm
	1 InGaAs (SWIR 1)	32 channels; 1.5–1.8 μm
	1 InSb (SWIR 2)	32 channels; 2.0–2.5 μm
	1 HgCdTe (LWIR)	12 channels; 8–12.5 μm
	1 InSb (MWIR)	3–5 μm (option)
IR detector cooling	Liquid nitrogen (closed cycle – option)	
IFOV	Choice of 1.25 mrad, 2.5 mrad, 3.3 mrad, or 5.0 mrad	
Swath angle	up to 90°	
Scan speed	up to 25 Hz with all bands, continuously selectable	
Pixels per line	512	
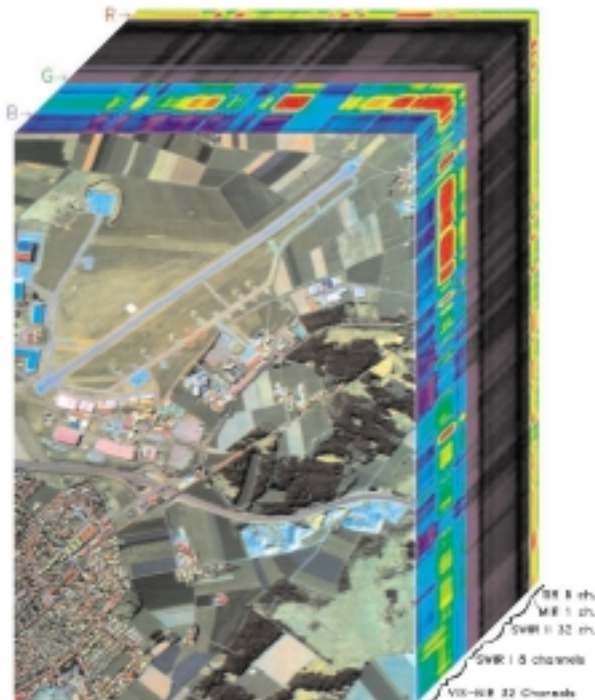
Fig. 28. Colour composite image recorded using the digital airborne imaging spectroradiometer (permission from German Aerospace Centre DLR).

The above-described system enables measurement of flux from a single spot in a number of spectral bands. However, because these systems employ a scanning system (typically the Kennedy-type reflective scanner) and are used in airborne applications, the imaging spectroradiometers generate two-dimensional image of the land below the aircraft in different spectral bands (see Fig. 28).

6. Light detection and ranging (LIDAR)

LIDAR is an acronym for light detection and ranging, a technique that uses laser light pulses to detect contaminations or aerosol in much the same way that sonar uses sound pulses, or radar uses radio waves. In radar, radio waves are transmitted into the atmosphere, which scatters some of the power back to the radar's receiver. A LIDAR also transmits and receives electromagnetic radiation, but at higher frequency. LIDARs operate in the ultraviolet, visible and IR region of the electromagnetic spectrum.

Different types of physical processes in the atmosphere are related to different types of light scattering. Choosing different types of scattering processes allows atmospheric composition, temperature and wind to be measured. The scattering is essentially caused by Rayleigh scattering on Nitrogen and Oxygen molecules, and Mie scattering on aerosols (dusts, water droplets, etc.). At low attitudes, Mie scattering is predominant because of the higher cross-section and the high aerosols concentration.

In the LIDAR approach, a laser radiation is transmitted into the atmosphere and backscattered radiation is detected as a function of time by optical receiver. The return time of the reflected or scattered pulses provides range information. In a LIDAR arrangement, the backscattered light is collected by a telescope, usually placed coaxially with the laser emitter. The signal is then focussed on a photodetector through a spectral filter, adapted to the laser wavelength. Different kinds of lasers are used depending on the power and wavelength required. The lasers may be both cw or pulsed. LIDARs typically use extremely sensitive detectors, which convert the individual quanta of light first into electric currents and then into digital photocounts, which next can be stored and processed on a computer.

In general, a signal is produced by direct absorption, fluorescence, or Raman scattering. Absorption techniques are most straightforward and widely applied. In the atmosphere, for example, long-path absorption spectroscopy is used in two wavelength bands – the infrared, where many molecules have characteristic fingerprints, and the ultraviolet (UV) to visible range.

In a typical case, the laser is alternatively tuned to a wavelength within the absorption band of interest (at λ_{on}) and then to a wavelength with negligible absorption (at λ_{on}), so that difference in the signal returned either from a surface or from air- or water-borne particles is recorded. By dividing the two LIDAR signals by each other, most troublesome and unknown parameters are eliminated and the gas concentration as a function of the range along beam can be evaluated. Such applications require tunable lasers, either tunable diode lasers in the IR or Nd:YAG dye lasers in the UV to visible range.

The principle for different absorption LIDAR (DIAL) is schematically represented in Fig. 29. Let us now assume that wavelength couple (λ_{on}, λ_{off}) is sent simultaneously into the atmosphere. As λ_{on} and λ_{off} have been chosen close enough for exhibiting the same scattering properties, the first chimney plume will cause an increase in the backscattering signal, because the concentration of aerosols is larger, but the same increase for both pulses. Conversely, the second chimney plume, which contains a certain quantity of the pollutant, will absorb the backscattered signal at the λ_{on}-wavelength much stronger than at the λ_{off}-one. From this difference, and using Beer-Lambert's law, one can deduce the specific concentration of the pollutant under investigation versus range. For typical pollutants, such as sulphur dioxide, nitrous, oxide, ozone, and mercury, detection ranges for the part-per-billion detection level are between 0.5 and 5 km.

The main alternative to direct-absorption spectroscopy is Raman scattering. This occurs when photons are non-elastically scattered from molecules, exciting them in the process and releasing some photon energy. Thus, Raman return signals are at a different, longer wavelength than the exciting wavelength. Raman cross-sections are
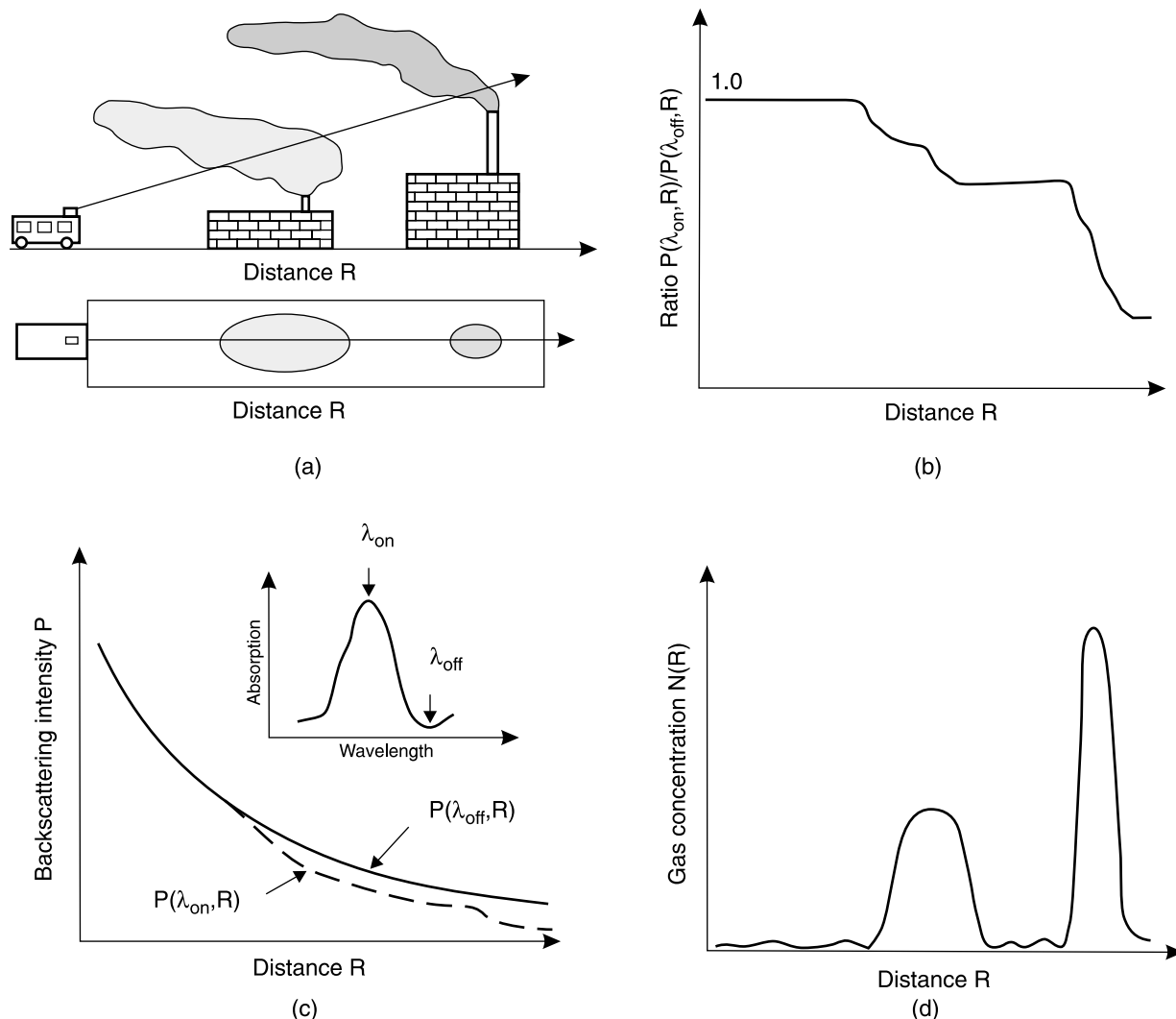
Fig. 29. Illustration of the principle of different absorption LIDAR (DIAL): (a) pollution measurement situation, (b) back-scattered laser intensity for the on- and off- resonance wavelengths, (c) ratio (DIAL) curve, and (d) evaluated gas concentration (after Ref. 20).

much smaller than absorption cross sections, so the Raman technique works well using high-power lasers only for higher concentrations (hundreds of parts per million) and distances of less than 1 km. Water vapour profiles can be obtained in vertical soundings up to several km in height, and pressure profiles up to tens of km are measurable using Raman signals from atmospheric N_2.

The other major technique, fluorescence spectroscopy, has limited use in atmospheric measurement because the return signal is too week. The technique is, however, an excellent way to monitor solid targets in the biosphere, such as oil, spills, algae bloom patches, and forest area. The fluorescent signal from plants originates in the excitation of chlorophyll and other leaf pigments. Fluorescence LIDAR is also a powerful technique for measurements at mesospheric heights where the pressure is low and the fluorescence is not quenched by collisions. This technique has been used extensively to monitor layers of various alkali and alkaline earth atoms (Li, Na, K, Ca, and Ca^+) at a height of about 100 km [20,21].

In addition to monitoring pollutants, LIDAR is widely used to measure wind velocities via Doppler shifts. Recently, improved laser stability has expended LIDAR to more ambitious projects, including the study of winds in the stratosphere.

The main advantage of LIDAR is that it can map the location of chemicals over a wide region. Due to the rapid nature of laser pulses, the time resolution is very critical (a few nanoseconds) to get good spatial resolution. Overall, DIAL systems can provide 2-D or 3-D information of air pollutants. However, most of the existing LIDAR systems have not met the pragmatic deployment requirements of users in industry or government. LIDAR systems are usually complex, large, expensive, and require high-skilled personnel for their operation.

7. Infrared gas sensors

Infrared gas detection is a well-developed measurement technology. In general, for toxic and combustible gas mon-

itoring applications, IR instruments are among the most users friendly and require the least amount of maintenance.

There are a number of ways by which various IR components can be arranged to produce a gas analyser. The design may be relatively simple, or very complicated depending on the type of analysers for the applications. Figure 30 illustrates some of the basic features of an IR analyser.
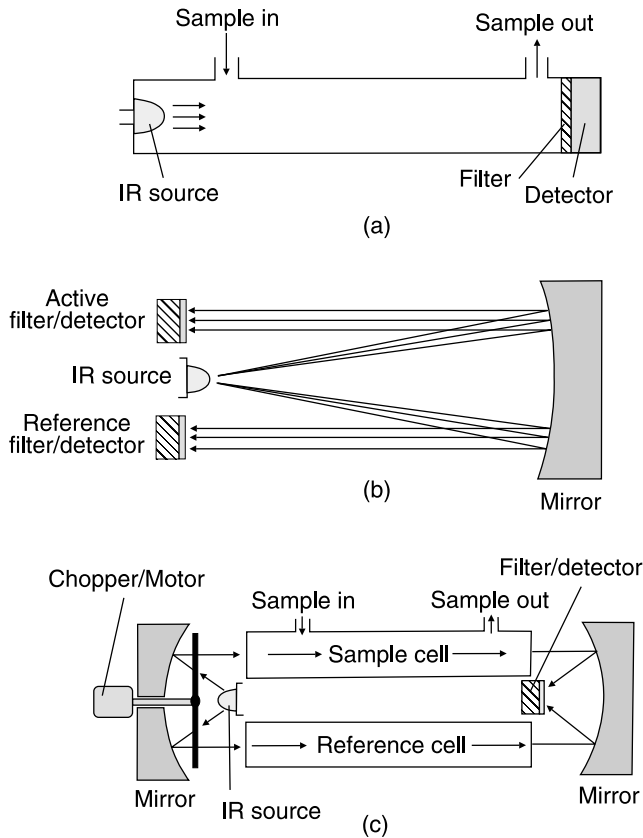

Fig. 30. Configurations of the gas analysers: (a) a basic gas detector layout, (b) a two-detector layout, and (c) double beams with chopper layout. Reproduced from Chou (2000).

The basic design is shown in Fig. 30(a), which consists of an IR source, bandpass filter and the interaction with the gas sample and detector. A detector is selectively sensitised to the absorption wavelength of the gas whose presence is to be detected by the use of a narrow bandpass optical filter. Clearly then, increased gas concentration in the optical path between source and filtered detector leads to a depression in signal level. The bandpass filter could be placed in front of the light source, instead of placing it in front of the detector.

In practice, in order to reduce false alarms and introduce a level of quantification, it is necessary to provide some calibration. Dependent on the application and instrument manufacturer, this may be taking the form of reference chamber containing a known concentration of the gas, or measurement of a reference wavelength just slightly outside the absorption band and/or dual matched detectors.

Figure 30(b) shows another popular design with layout of two detectors. Modulated flashing IR sources are reflected back to the detectors. In this arrangement, the active detector has a filter for the target gas, while the reference detector has a filter with different wavelength. In such a way, the active detector is used to detect the target gas and the reference detector is used to ignore the target gas. In actual operation, the reference detector provides a base point value (or zero point) while the active detector is used to provide the signal. An advantage of this design is compensation of changes that occur in the detector's sensitivity with time (for example, change in the intensity of the light source).

The design illustrated in Fig. 30(c) uses two tubes or cells. One is a reference cell that is filled with a pure target or reference gas, while the other is a sampling cell in which the sample gas passes through. Additionally, a chopper in the form of disc with a number of slots in it is used. As the chopper rotates, it alternately allows the light beam to pass through the sample and reference cells. The detector gets its base reading from the reference cell.

There are many light sources available, ranging from regular incandescent light bulb to specially designated heating filaments and electronically generated sources. The last sources are used to generate enough radiation at the wavelength of interest for the purpose of detecting the specific target gas. A heated wire filament, similar to that in a pen flashlight, is used in the 1–5-μm spectral range for the detection of most hydrocarbons, carbon dioxide, and carbon monoxide. Alternatives include glowbars (rods of silicon carbide) or coils, typically of nichrome alloy resistance wire with high emissivity in the MWIR region.

Figure 31 shows the spectral response of lead selenide photoconductors superimposed over the absorption bands of the common air pollutants: HC, SO_2, CO, and NO. Such
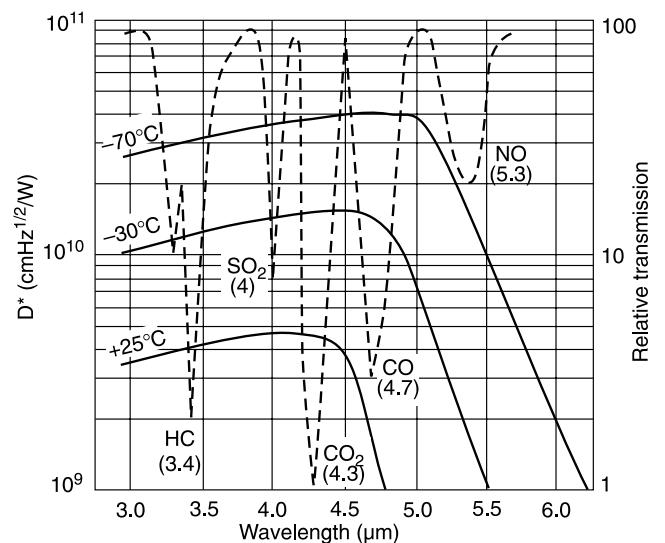

Fig. 31. Detectivity of lead selenide detectors as a function of wavelength at operating temperatures of +25°C, –30°C and –70°C, along with absorption spectra of same common air polluters.

detectors are optimally matched to the 3–5-μm range. Beyond that wavelengths, the designer tends to be constrained to thermal detectors, the increased cost of the only suitable photon detector (HgCdTe) outweighing the performance benefits in most cases. Hence the compromise is stated; lower performance over a very broad waveband or higher performance in a more limited one. For most specific applications, the more limited wavebands do contain sufficient structure and thus photon detectors are preferred. On the cost side, thermopiles are the cheapest option, and may be favoured when lower performance is acceptable. In most cases, however, the significant performance benefits of lead salt detectors more than compensate for their relatively small additional cost.

Infrared detection is applied to numerous applications. The most important are monitoring: in transport industry (car and truck exhausts), petrochemical industry (gas leaks in refineries and oil rigs), and medicine (carbon dioxide exhalation and anaesthetic gases).

References

1. W. Herschel, "Experiments on the refrangibility of the invisible rays of the Sun," *Phil. Trans. Roy. Soc. London* **90**, 284 (1800).
2. P.R. Norton, "Infrared detectors in the next millennium," *Proc. SPIE* **3698**, 652–665 (1999).
3. S.G. Burnay, T.L. Williams, and C.H. Jones, *Applications of Thermal Imaging*, Adam Hilger, Bristol, 1988.
4. R.D. Hudson, *Infrared System Engineering*, Wiley, New York, 1969.
5. A. Rogalski, *Infrared Detectors*, Gordon and Breach Science Publishers, Amsterdam, 2000.
6. J.L. Miller, *Principles of Infrared Technology*, Van Nostrand Reinhold, New York, 1994.
7. M.E. Couture, "Challenges in IR optics," *Proc. SPIE* **4369**, 649–661 (2001).
8. D.C. Harris, *Materials for Infrared Windows and Domes*, SPIE Optical Engineering Press, Bellingham, 1999.
9. W.J. Smith, *Modern Optical Engineering*, McGraw-Hill, New York, 2000.
10. J.M. Lloyd, *Thermal Imaging Systems*, Plenum Press, New York, 1975.
11. A.A. Cameron, "The development of the combiner eyepiece night vision goggle," *Proc. SPIE* **1290**, 16–19 (1990).
12. I.P Csorba, *Image Tubes*, Howard Sams, Indianapolis, 1985.
13. M. Hewish, "Night-vision goggles," *Defense Electronics & Computing*, Supplement to *IDR* **2**, 17–24 (1992).
14. S.B. Campana (editor), *The Infrared and Electro-Optical Systems Handbook,* Vol. 5, *Passive Electro-Optical Systems*, SPIE Optical Engineering Press, Bellingham, 1993.
15. L.J. Kozlowski and W.F. Kosonocky, "Infrared detector arrays," in *Handbook of Optics,* Chap. 23, edited by M. Van Bass, E.W. Stryland, D.R. Williams, and W.L. Wolfe, McGraw-Hill Inc., New York, 1995.
16. L.J. Kozlowski, K. Vural, J. Luo, A. Tomasint, T. Liu, and W.E. Kleinhans, "Low-noise infrared and visible focal plane arrays," *Opto-Electron. Rev* **4**, 259–269 (1999).
17. STANAG No. 4349, Measurement of the minimum resolvable temperature difference (MRTD) of thermal cameras.
18. M. Hewish, "Lifting the veil of darkness," *Jane's International Defence Review* **6**, 49–56 (1997).
19. K. Chrzanowski, *Non-contact Thermometry-Measurement Errors*, Research and Development Treaties, Vol. 7, SPIE Polish Chapter, Warsaw, 2001.
20. S. Svanberg, "Environmental monitoring using optical techniques, in *Applied Laser Spectroscopy*, pp. 417–434, edited by W. Demtröder and M. Inguscio, Plenum Press, New York, 1990
21. J.P. Wolf, H.J. Kölsch, P. Rairoux, and L. Wöste, "Remote detection of atmospheric pollutants using differential absorption lidar techniques," in *Applied Laser Spectroscopy*, pp. 435–467, edited by W. Demtröder and M. Inguscio, Plenum Press, New York, 1990.
22. J. Chou, *Hazardous Gas Monitors*, McGraw-Hill Book Company, New York, 2000.